Astronomy Reports, Vol. 44, No. 10, 2000, pp. 641–646. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 723–729.

The Active Galaxy 3C 66A: A Variable Source of Very High-Energy Gamma-Rays

Yu. I. Neshpor, A. A. Stepanyan, O. R. Kalekin, N. A. Zhogolev, V. P. Fomin, N. N. Chalenko, and V. G. Shitov

Crimean Astrophysical Observatory, p/o Nauchnyĭ, Crimea, 334413 Ukraine

Received September 20, 1999

Abstract—Observations of the very-high-energy gamma-ray flux of the blazar 3C 66A ($z = 0.444$) carried out at the Crimean Astrophysical Observatory with the GT-48 atmospheric Cerenkov detector are reported. The gamma-ray fluxes in 1997 and 1998 were lower than in 1996. The optical luminosity of the object in 1997–1998 also decreased in comparison with its value in 1996. If the emission is isotropic, the very-high-energy gamma-ray power is 10^{46} erg/s. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Blazars—active galactic nuclei whose prototype is BL Lac—are probable extragalactic sources of gamma rays with very high energies (VHE, $E > 10^{11}$ eV). These objects are very interesting astrophysically, because they are characterized by substantial flux variations at all wavelengths from radio to X-rays. In some cases, large-amplitude X-ray variations coincide with optical variations [1]. The timescales of the variations are from minutes to about one year. Blazars are distinguished by a strong tendency for flaring (with timescales of several days) and outburst activity (with durations of several months). This suggests that these objects contain a large number of high-energy particles, which can generate VHE gamma rays when they interact with matter or electromagnetic fields.

The first such objects from which VHE gamma rays were detected were the BL Lac objects Mrk 421 and Mrk 501 [2, 3]. The related active galaxy 3C 66A was detected optically [4] as a 15th magnitude pointlike blue source. Its optical emission is strongly polarized, with the degree of polarization varying widely with time, sometimes reaching 30%. The brightness of the object also varies appreciably with time; for instance, it exceeded 14^m in 1996.

High-energy gamma rays were detected toward 3C 66A at energies > 100 MeV by EGRET, and the gamma-ray source was given the name 2EG J0220+4228. The position error of the EGRET measurement was 1° [6]. Additional measurements of the gamma-ray flux from this direction allowed the source position to be determined to higher accuracy. In the third EGRET catalog [7], this object was listed as the gamma-ray source 3EG J0222+4253 with coordinates $\alpha = 35\overset{\circ}{.}7$ and $\delta = 42\overset{\circ}{.}9$, which differ from the coordinates of 3C 66A by $0\overset{\circ}{.}15$.

We detected VHE gamma rays from the blazar 3C 66A in 1996 in our observations on the GT-48 Cerenkov detector [8]. The observations were continued in 1997 and 1998. The results of the three-year flux measurements are presented below.

2. BRIEF DESCRIPTION OF THE GT-48 GAMMA-RAY TELESCOPE

Gamma rays with energies $E > 10^{11}$ eV can be detected using ground-based equipment, using the fact that VHE gamma rays interact with the nuclei of atoms in the air, generating so-called electron–photon showers. These are made up of high-energy electrons and positrons, which emit optical Cerenkov radiation, primarily at small angles ($0\overset{\circ}{.}5$–1°) to the direction of motion of the primary photon. This makes it possible to determine the direction from which the flux of gamma rays arrives.

The effective area in which Cerenkov events from electron–photon showers can be detected is rather large. The region on the Earth covered by an event with vertical incidence of the primary particle is approximately 4×10^8 cm^2; i.e., it forms a circle with a diameter of 250 m. This makes it possible to detect small (approximately 10^{-11} photons cm^{-2} s^{-1}) gamma-ray fluxes.

The main obstacle to detecting and studying VHE gamma-ray sources is the presence of an appreciable cosmic-ray background, whose particles cause Cerenkov events in the Earth's atmosphere that are difficult to distinguish from those due to discrete gamma-ray sources. Nevertheless, there are some differences between these two types of events. Currently, multielement receiving cameras are used to make images of Cerenkov events. Differences in the parameters of the images of Cerenkov events resulting from gamma rays and from cosmic-ray particles enable us to eliminate most of the latter events.

1063-7729/00/4410-0641$20.00 © 2000 MAIK "Nauka/Interperiodica"

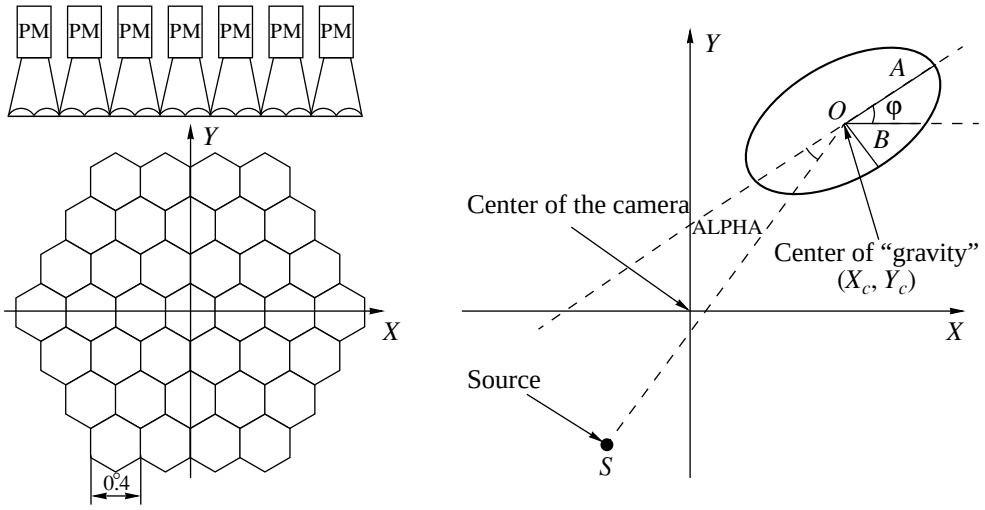


Fig. 1. Schematic representation of the light detector and parameters of a Cerenkov event. *A* is the effective length and *B* the effective width of the event image, φ is the orientation angle, ALPHA is the azimuthal angle, and PM is the photomultiplier.

The first telescope with a multichannel camera began operation at the Whipple Observatory (USA) in 1982 [9]. A similar telescope, GT-48, started working at the Crimean Astrophysical Observatory in 1989. We have described the GT-48 telescope in a number of papers (see, for example, [10]). The facility consists of two identical northern (1) and southern (2) altitude–azimuth mounts (sections) separated by 20 m in the north–south direction at a height of 600 m above sea level. We showed in [11] that, in contrast to single telescopes, a double telescope operating in a coincidence regime can almost completely eliminate events due to individual charged particles hitting the light detectors.

Six aligned telescopes are mounted on each section. The optics of each telescope consist of four 1.2-meter mirrors with a common focal point. The mirrors of three telescopes have a focal length of 5 m. Light detectors (cameras) consisting of 37 photomultipliers (37 cells) that can image Cerenkov events at visual wavelengths (300–600 nm) are located in their focal plane. Events are recorded only when the amplitudes of time-coincident signals in any two of the 37 cells exceed a preset threshold. The time resolution of the coincidence circuit is 15 ns.

There is a conical hexagon-shaped light-guide in front of each photomultiplier. The mean diameter of the entrance window corresponds to the linear angle of the field of view of one cell, $0.^{\circ}4$ (Fig. 1). The field of view of the entire light detector is $2.^{\circ}6$.

The other three telescopes have focal lengths of 3.2 m and are intended for detection of the ultraviolet radiation of Cerenkov events at 200–300 nm. The detectors are sun-blind photomultipliers.

The total area of the mirrors on both mountings (sections) is 54 m^2. The installation can be moved by a control system with a drive accuracy of $\pm 1'$. Observations can be carried out both in a coincidence regime using the two sections and independently by each section. The effective threshold energy for detection of gamma rays is 1.0 TeV.

3. OBSERVATIONS AND DATA PROCESSING

Observations of 3C 66A ($\alpha = 2^{\circ}22'40''$, $\delta = 43^{\circ}02'08''$) were carried out in 1996, 1997, and 1998 using the two sections in a coincidence regime with a time resolution of 100 ns. We tracked the object and an area of sky (background) at the same azimuth and zenith angles with a 30-min time shift between trackings (the duration of a single observation was 25 min). The observations of the background preceded those of the source.

In total, we processed the data for 12 sessions in 1996, 30 in 1997, and 17 in 1998, corresponding to a total duration for the observations of 3C 66A of 1175 min (24 h 35 min). We performed a preliminary reduction of the data, necessary for correct calculation of the first and second moments of the brightness distribution, from which we derived the parameters of the Cerenkov events: effective length *A*, effective width *B*, angle φ describing the direction of maximum elongation of the image (i.e., its orientation), and the coordinates X_c and Y_c of the center of "gravity" of the brightness distributions (Fig. 1). We calculated the moments for cells whose signal exceeded a certain threshold value [12]. All other parameters of a Cerenkov event can be derived from these parameters [10].

As a result of our preliminary reduction, there remained over the three years of observations 34695 source events and 34770 background events for further analysis. Thus, the difference of the number of source (N_s) and background (N_b) events $N_\gamma = N_s - N_b = -75 \pm 264$, where 264 is the statistical error. This includes a contri-

bution from transparency variations. To determine the probable gamma-ray flux, we must eliminate events due to the charged cosmic-ray component.

We noted above that the parameters of Cerenkov events due to VHE gamma rays differ little from those due to charged cosmic-ray particles. Nevertheless, by excluding events due primarily to cosmic rays, we can considerably reduce the error in the difference in the number of source and background events. This requires correct choice of the boundary values of the selection parameters, in order to optimize the signal-to-noise ratio $Q = (N_{0_s} - N_{0_b})/\sqrt{N_{0_s} + N_{0_b}}$, where N_{0_s} and N_{0_b} are the numbers of gamma-like events in the on-source and background observations. The difference $N_{0_s} - N_{0_b} = N_\gamma$ is the selected number of gamma rays detected during the observations, and $\sqrt{N_{0_s} + N_{0_b}}$ is the statistical error of the signal determination after selection. If the selection is done using several parameters, it is possible to exclude up to 99% or more of events due to the charged cosmic-ray component.

When analyzing the 1996 data in [8], we used the parameter DRO, which characterizes the magnitude of the stereoscopic effect. Selecting according to this parameter, we could detect the flux of VHE gamma rays from 3C 66A at the 5.1σ level. However, our analysis of observations for the powerful VHE gamma-ray source Mrk 501 [13], detected at the 11σ level, indicated that the accuracy of the estimated position of a VHE gamma-ray source decreases when DRO is used. Though using DRO enables us to efficiently select gamma showers against the background of showers from charged particles, the directions of the axes of the selected showers have relatively large errors. Therefore, we decided to repeat the analysis of the 1996 observations without using DRO.

The parameters of events registered simultaneously by each section were determined independently using the data for each section, so that each event had two values for each parameter, which we will label "1" for the northern section and "2" for the southern section. To reduce the background due to cosmic-ray particles, we applied a number of selection criteria based on the parameters of the Cerenkov events (Fig. 2).

In the selection, we considered first and foremost the total "energy" of the event V (the integrated light flux), which was measured over the same area as the second moments of the events in units of the analog-code converter quantization step. Events with amplitudes $V(1) < 75$ units or $V(2) < 125$ units were excluded from further consideration. Other parameters used to distinguish gamma-ray showers against the background of showers from charged particles (p showers) were the effective length and width of the image of the event. We excluded events from further consideration if at least one of the following conditions was fulfilled: $A(1) > 0.^\circ33$, $A(2) > 0.^\circ33$, $B(1) > 0.^\circ23$, or $B(2) > 0.^\circ23$.

As noted above, we also recorded the ultraviolet radiation of the events. The electrons from p showers with a given energy are known to penetrate, on average, to appreciably greater depths in the Earth's atmosphere than those from gamma-ray showers with the same energy. Thus, Cerenkov events from p showers have considerably greater radiation fluxes at 200–300 nm (in the ultraviolet) [14]. We will call the ratio of an event amplitude at these wavelengths (U) to the total amplitude in the visual (V) the parameter UV. This parameter was first successfully used by us in our analysis of observations of the Crab Nebula [15].

On average, over all the years of observation, selection according to UV has improved the confidence of the detections up to the 5.9σ level. Use of this parameter increases the signal-to-noise ratio by a factor of 2.5. The parameters A, B, V, and UV do not depend on the location of the event relative to the source, and are coordinate-independent.

Application of coordinate-dependent parameters (such as, for example, ALPHA; Fig. 1) enables us to increase Q (the signal-to-noise ratio) and improve our estimate of the direction toward the gamma-ray source. In this case, selection is performed for each section separately. Therefore, the set of events selected for one section only partially overlaps with the set for the other section. We can analyze such sets of gamma-like events in order to improve the coordinates of the gamma-ray source and more accurately determine the magnitude of the gamma-ray flux.

4. RESULTS

To determine the direction of the incoming gamma-ray flux, we used a trial-source method [16–18]. This method is based on the fact that, in the focal plane of the telescope, the images of gamma-ray events are oriented toward the source, whereas the major axes of the image ellipses of p showers are oriented to first approximation uniformly in all directions. Let us suppose that we first select events adopting for the source direction an arbitrary point in the focal plane with coordinates X_i and Y_j, then make a selection using coordinate-dependent parameters. In this case, the number of remaining p showers will not depend on the adopted source position. At the same time, the number of images from gamma-ray showers will strongly depend on the adopted source position and will have a maximum in the direction toward the true source. In our case, the direction toward the source coincided with the center of the camera.

We can plot the distribution of the number of events selected over the field of view of the light detector as a function of the adopted source position $N(X_i, Y_j)$. This is essentially a "map", or three-dimensional histogram, in which two dimensions are the Cartesian coordinates of the trial source relative to the center of the detector field of view, and the third dimension is the number of gamma-like events selected using a coordinate-depen-

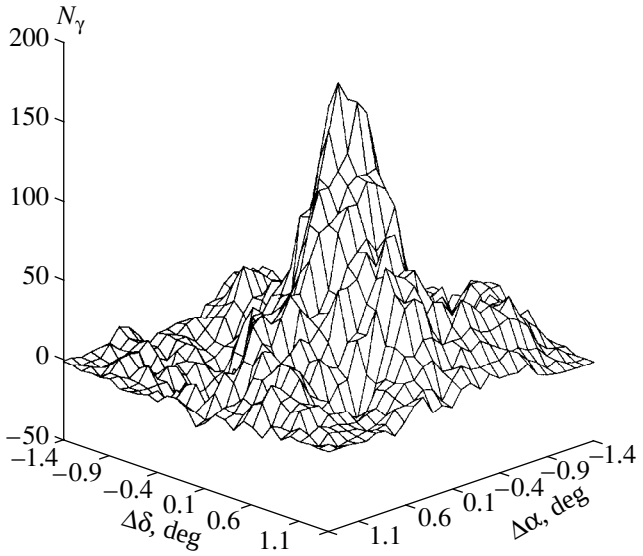


Fig. 2. Three-dimensional histogram for the selected gamma-ray events from 3C 66A. Δδ is the deviation from the position of 3C 66A in declination, and Δα is the deviation in right ascension (in deg).

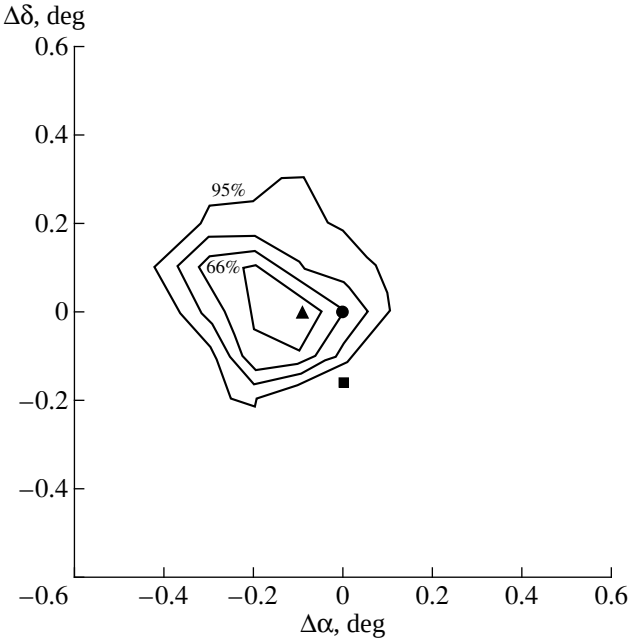


Fig. 3. Contours of N_γ for the gamma-ray source 3C 66A. Circle: position of the blazar 3C 66A; triangle: VHE gamma-ray source; square: gamma-ray source 3EG J0222 + 4253. Δδ is the deviation from the position of 3C 66A in declination, and Δα is the deviation in right ascension (in deg).

dent parameter. In this way, we can find the true position of the gamma-ray source. We have plotted such histograms for the data from both the northern and southern sections. We carried out the selection using both the coordinate-independent parameters noted above and the parameters DIST and MISS. DIST is numerically equal to the angular distance from the event center of "gravity" to the trial source, and MISS = DIST sin(ALPHA) (Fig. 1). We set the limits on the values of these parameters DIST(1) < 0°.9 and 0°.25 < DIST(2) < 0°.95. We excluded events with MISS(1) > 0°.210 and MISS(2) > 0°.225 for sections 1 and 2, respectively.

We can construct a total map from the maps for the two sections. Correct determination of the statistical error requires that events coincident at both sections be counted as a single event. Therefore, we found for each section gamma-like events $N_c(X_i, Y_j)$ present in both sections, and also gamma-like events that, after selection in coordinate-dependent parameters, remained in the data for only the northern ($N_1(X_i, Y_j)$) or southern ($N_2(X_i, Y_j)$) section. In this case, the number of registered events identified as gamma rays in the on-source observations will be $N_s(X_i, Y_j) = N_{c_s}(X_i, Y_j) + N_{1_s}(X_i, Y_j) + N_{2_s}(X_i, Y_j)$. Similarly, this quantity for the background data will be $N_b(X_i, Y_j) = N_{c_b}(X_i, Y_j) + N_{1_b}(X_i, Y_j) + N_{2_b}(X_i, Y_j)$. We must map the background in order to exclude instrumental and methodical effects (see, for example, [17, 19]). The difference $N_s(X_i, Y_j) - N_b(X_i, Y_j) = N_\gamma(X_i, Y_j)$ enables us to determine the coordinates of the gamma-ray source to within a few tenths of a degree. The position of the maximum of N_γ on the map corresponds to the direction toward the observed gamma-ray source. We can write the statistical error of N_γ

$$\sigma = \sqrt{(N_{c_s} + N_{1_s} + N_{2_s}) + (N_{c_b} + N_{1_b} + N_{2_b})}.$$

Figure 2 presents the resulting three-dimensional histogram for the MISS criterion. The maximum value of $N_\gamma = 172 \pm 29$ (corresponding to a 5.9σ detection) has coordinates $X_i = -0°.1$ and $Y_j = 0°.0$. During the observations, the center of the camera was pointed toward the blazar 3C 66A. Figure 3 shows the isophotes of N_γ for this histogram. The initial value of the isophote is drawn at the 95% significance level; this corresponds to an error box of 0°.50. The most probable location of the high-energy gamma-ray source 3EG J0222 + 4253 detected by EGRET [7] (the error box at the 95% significance level is 0°.31) and the position of 3C 66A are plotted.

Visual observations [20, 21] show that, in October–November 1997, the brightness of 3C 66A decreased by approximately one magnitude compared to the same period in 1996, and a minor brightness increase was again observed in 1998. Analysis of our VHE observations (after selection of gamma-like events) shows that the gamma count rate was $N_{96} = (0.290 \pm 0.053)$ min^{-1} in 1996, $N_{97} = (0.060 \pm 0.023)$ min^{-1} in 1997, and $N_{98} = (0.094 \pm 0.039)$ min^{-1} in 1998. This implies that the gamma-ray flux from the direction toward 3C 66A appreciably decreased in 1997 and 1998 compared to 1996. A comparison of modeling results with the observa-

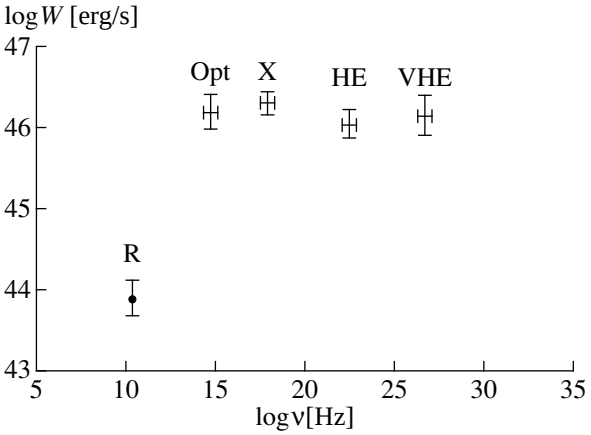


Fig. 4. Spectrum of the galaxy 3C 66A. The vertical segments indicate the intervals for flux variations. The horizontal axis plots the logarithm of the frequency, and the vertical axis plots the logarithm of the power per logarithmic frequency interval (νF). R denotes radio emission, Opt optical emission, X X-ray emission, HE gamma rays with energies above 100 MeV, and VHE gamma rays with energies above 1 TeV.

tional data shows that, on average for the three years of observations, the flux of gamma rays with energies >1 TeV was $(3.0 \pm 0.9) \times 10^{-11}$ photons cm^{-2} s^{-1}.

5. CONCLUSIONS

Our multivariate analysis (using multiple parameters describing the Cerenkov events, both coordinate-independent and coordinate-dependent) enables us to state with a high degree if confidence (at about the 6σ level) that there exists a VHE gamma-ray source ($E_{\rm thr} >$ 1.0 TeV) toward the galaxy 3C 66A. The good coincidence of the gamma-ray source position derived using a trial-source method with that of the blazar 3C 66A further increases the trustworthiness of the results obtained. In 1997, we observed a decrease of the VHE gamma-ray flux from the object compared to its 1996 flux, simultaneous with a decrease in the visual brightness of 3C 66A; this may provide additional evidence that we have detected gamma-ray emission precisely from this source.

We note especially that, whereas the position of the high-energy gamma-ray source 2J 0220+4228 in the second EGRET catalog [6] differed from that of the VHE gamma-ray source detected by the GT-48 gamma-ray telescope by 1°, additional higher accuracy EGRET observations [7] yielded a position for the high-energy source 3J 0222 + 4253 that agrees within $0\overset{\circ}{.}15$ with the position of the VHE gamma-ray source (Fig. 3) and the blazar 3C 66A. According to the EGRET data averaged over five years of observations, the flux of gamma rays with energies above 100 MeV is $(18.7 \pm 2.9) \times 10^{-8}$ photons cm^{-2} s^{-1}. The differential spectral index is 2.01 ± 0.14. According to our data, the flux of gamma rays with energies above 1 TeV is $(3.0 \pm 0.9) \times 10^{-11}$ photons cm^{-2} s^{-1}. If we suppose that the spectral index does not change in this energy range, it is equal to –1.95, in good agreement with the data at 100 MeV. Of course, this comparison is not completely correct, since the gamma-ray flux from 3C 66A is variable at both 1 TeV and 100 MeV. Nevertheless, this result is of some interest.

According to Lanzetti [22], the redshift of 3C 66A is $z = 0.444$. If we assume that the gamma rays are emitted isotropically, their power is approximately 10^{46} erg/s. On the one hand, this power may be strongly overestimated, since the VHE gamma-ray radiation is probably not isotropic, but it is rather difficult to estimate the degree of anisotropy of the radiation. On the other hand, at such great distances, the flux of gamma rays with energies >1 TeV is strongly attenuated due to interactions with optical radiation in intergalactic space. According to the estimates of Stecker and de Jager [23], the flux of photons from 3C 66A with energies >1 TeV is attenuated by a factor of about 200. Therefore, at present, it is not possible to provide a firm estimate of the source's true VHE power.

It is interesting to compare the power of the VHE emission of 3C 66A with its power at other frequencies. Figure 4 presents the power per logarithmic frequency band νF in the radio, optical, and X-ray, as well as the high-energy and VHE gamma-ray ranges [1, 7, 20]. It is striking that the power in all ranges except for the radio is approximately 10^{46} erg/s. Note that the spectrum of 3C 66A is rather similar to that of the well-studied galaxy Mrk 501 (see, e.g., [24]), but its absolute power is two orders of magnitude higher. The similarity of the spectra suggests a common nature for these two objects and their radiation mechanisms.

ACKNOWLEDGMENTS

The authors are grateful to the employees of the Crimean Astrophysical Observatory Z.N. Skiruta and S.G. Kochetkova for processing the observational data.

REFERENCES

1. D. Maccagni, B. Garilli, R. Schild, and M. Tarenghi, Astron. Astrophys. **178**, 21 (1987).
2. M. Punch, C. W. Akerlof, M. F. Cawley, *et al.*, Nature **358** (6386), 477 (1992).
3. M. Catanese, C. W. Akerlof, S. Biller, *et al.*, in *Padova Workshop on TeV Gamma-Ray Astrophysics*, Ed. by M. Cresti, p. 348.
4. B. J. Wills and D. Wills, Astrophys. J. **190**, L97 (1974).
5. Yu. S. Efimov and N. M. Shakovskoy, in *Proceedings of the OJ-94 Annual Meeting, Perugia, Italie, Sept. 8–9, 1997*, Ed. by G. Tosti and L. Takalo, p. 24.
6. D. J. Thompson, D. L. Bertsch, B. L. Dingus, *et al.*, Astrophys. J., Suppl. Ser. **101**, 259 (1995).
7. R. C. Hartman, D. L. Bertsch, S. D. Bloom, *et al.*, Astrophys. J., Suppl. Ser. **123**, 79 (1999).

8. Yu. I. Neshpor, A. A. Stepanyan, O. R. Kalekin, *et al.*, Pis'ma Astron. Zh. **24**, 167 (1998) [Astron. Lett. **24**, 134 (1998)].
9. M. F. Cawley, J. Clear, D. J. Fegan, *et al.*, in *Proceedings of the 18th International Cosmic Ray Conference, Bangalore, India, 1983*, Vol. 1, p. 118.
10. B. M. Vladimirskiĭ, Yu. L. Zyskin, A. A. Kornyenko, *et al.*, Izv. Krym. Astrofiz. Obs. **91**, 74 (1995).
11. N. N. Chalenko, O. R. Kalekin, Yu. I. Neshpor, and A. A. Stepanyan, J. Astrophys. Astron. **18**, 151 (1997).
12. A. P. Kornienko, A. A. Stepanyan, and Yu. L. Zyskin, Astropart. Phys. **1**, 245 (1993).
13. O. R. Kalekin, N. N. Chalenko, Yu. L. Zyskin, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 604 (1999).
14. A. A. Stepanyan, V. P. Fomin, and B. M. Vladimirskiĭ, Izv. Krym. Astrofiz. Obs. **66**, 234 (1983).
15. O. R. Kalekin, Yu. I. Neshpor, A. A. Stepanyan, *et al.*, Pis'ma Astron. Zh. **21**, 184 (1995) [Astron. Lett. **21**, 163 (1995)].
16. C. W. Akerlof, M. F. Cawley, M. Chantell, *et al.*, Astrophys. J. Lett. **377**, L97 (1991).
17. Yu. I. Neshpor, A. P. Kornyenko, A. A. Stepanyan, and Yu. L. Zyskin, Exp. Astron. **5**, 405 (1994).
18. V. P. Fomin, S. Fennell, R. C. Lamb, *et al.*, Astropart. Phys. **2**, 151 (1994).
19. A. P. Kornyenko, Yu. I. Neshpor, Yu. L. Zyskin, and A. A. Stepanyan, Izv. Krym. Astrofiz. Obs. **93**, 143 (1996).
20. L. O. Takalo, T. Parsimo, A. Sillanpaa, *et al.*, http://bldata.pg.infn.it./volume_1/6/6.html (1999).
21. Yu. S. Efimov, personal communication (1999).
22. K. M. Lanzetta, D. A. Turnshek, and J. Sandoval, Astrophys. J., Suppl. Ser. **84**, 109 (1993).
23. F. W. Stecker and O. C. de Jager, in *Proceedings of the Kruger National Park Workshop on TeV Gamma-Ray Astrophysics*, Ed. by O. C. de Jager, p. 39.
24. J. Kataoka, J. R. Mattox, J. Quinn, *et al.*, Astropart. Phys. **11** (1–2), 149 (1999).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 647–653. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 730–736.

Amplification of Spiral Density Waves in Gaseous and Dust Galactic Disks as a Result of Interaction with the Halo

V. A. Antonov and A. S. Baranov

Pulkovo Astronomical Observatory, Russian Academy of Sciences, Pulkovskoe sh. 65, St. Petersburg, 196140 Russia

Received December 25, 1998

Abstract—The effect of the discrete structure of the halo on local oscillations of a galactic disk is analyzed. Such effects have much in common with dynamical friction. Gaseous and stellar disks are considered; in both cases, some leading spiral density waves are unstable. Bending oscillations of the disk can also be unstable when the disk interacts with the halo. © *2000 MAIK "Nauka/Interperiodica".*

1. FORMULATION OF THE PROBLEM

Studies of the stability of galactic disks and density-wave propagation traditionally treat galactic halos as a passive background creating a gravitational field. However, in any nonstationary process, the halo is distorted along with the disk, which should lead to a feedback effect on the disk. Such effects can accumulate substantially over time intervals comparable to the age of the galaxy. To some extent, these effects can resemble dynamical friction, with the density wave playing the role of the massive body. In this case, halo stars are assumed each time to approach the wave as if for the first time, without remembering previous interactions. This assumption seems to be quite plausible and consistent with the classical formulation of the dynamical friction problem [1–3]. In addition, regularly accumulating resonant interactions are possible between certain halo and disk zones. However, resonant interactions are a rather specific issue [4] and we do not address it in this paper. We will consider the gaseous and stellar disks separately.

2. A GASEOUS DISK

We begin with the gaseous disk and ignore effects due to its finite thickness, which are of little importance for the problem at hand. We thus use a thin planar gas approximation. We apply a local analysis and assume all conditions to be uniform in the coordinates x and y in the equatorial plane of the galaxy ($z = 0$). Provided the coordinate system orientation is chosen appropriately, a surface density perturbation $\sigma(x, y)$ can be written in the form

$$\delta\sigma = ae^{\lambda t + ikx} \quad (k > 0), \tag{1}$$

where λ and k are the usual increment and wave number, respectively, a is an amplitude, δ the linearized increment, and t time.

A density increase generates the corresponding potential perturbation $\delta\phi$, which we determine below. In contrast to usual density-wave theory, this perturbation plays a twofold role in our analysis: it affects both the gas itself and the halo stars. This results in a redistribution of the halo density, which, in turn, gives rise to a correction potential $\delta\phi_1$. At the end of the computations, we must include this correction in the equations for the gas-dynamical disk oscillations.

Since stars move in three dimensions, it is convenient to make a transformation from $\delta\sigma$ to the space density $\rho(x, y, z)$ using the Dirac δ function:

$$\begin{aligned}\delta\rho = \delta\sigma\bar{\delta}(z) &= \delta\sigma\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\mu z}d\mu \\ &= \frac{a}{2\pi}e^{\lambda t}\int_{-\infty}^{+\infty} e^{ikx + i\mu z}d\mu,\end{aligned} \tag{2}$$

where $\bar{\delta}$ is the delta function and μ the z component of the wave number vector.

The perturbation of the density $\exp(ikx + i\mu z)$ corresponds to that of the potential $-4\pi G[\exp(ikx + i\mu z)]/(k^2 + \mu^2)$, so that

$$\delta\phi = -2aGe^{\lambda t}\int_{-\infty}^{+\infty}\frac{e^{ikx + i\mu z}}{k^2 + \mu^2}d\mu, \tag{3}$$

where G is the gravitational constant. Integrating this expression over μ yields the usual relation between $\delta\sigma$ and $\delta\phi$ [5], however, the integral form (3) is more convenient in our case.

1063-7729/00/4410-0647$20.00

Stellar motions in the halo are described by the Boltzmann equation:

$$\frac{\partial f}{\partial t} + u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} - \frac{\partial \phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \phi}{\partial z}\frac{\partial f}{\partial w} = 0, \quad (4)$$

where f is the phase-space density and u, v, and w are the Cartesian components of the velocity vector $\mathbf{V}$. We then linearize Eq. (4) to obtain

$$\frac{\partial \delta f}{\partial t} + u\frac{\partial \delta f}{x} + v\frac{\partial \delta f}{\partial y} + w\frac{\partial \delta f}{\partial z} - \frac{\partial \delta\phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \delta\phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \delta\phi}{\partial z}\frac{\partial f}{\partial w} = 0 \quad (5)$$

(this approach is stretching things slightly, since linearization implies that we ignore terms with gradϕ components; i.e., we actually assume the unperturbed stellar motions in the halo to be inertial). Equation (5) can be solved for δf using an ordinary Fourier expansion. We thus obtain

$$\delta f = -2iaGe^{\lambda t} \times \int_{-\infty}^{+\infty} \frac{\left(k\frac{\partial f}{\partial u} + \mu\frac{\partial f}{\partial w}\right)e^{ikx + i\mu z}}{(k^2 + \mu^2)[\lambda + i(ku + \mu w)]}d\mu. \quad (6)$$

Strictly speaking, we should bear in mind the anisotropy of the velocity distribution. However, for general orientational results, we adopt a spherically symmetric velocity distribution for the halo stars: $f = F(V)$. It then follows that

$$\delta f = -\frac{2iaGe^{\lambda t}F'(V)}{V} \times \int_{-\infty}^{+\infty} \frac{(ku + \mu w)e^{ikx + i\mu z}}{[\lambda + i(ku + \mu w)](k^2 + \mu^2)}d\mu. \quad (7)$$

To transform to the mass density $\delta\rho_1$, we must integrate δf over velocity space and multiply by the stellar mass $\bar{m}$:

$$\delta\rho_1 = -2ia\bar{m}Ge^{\lambda t}\int\frac{F'(V)}{V}d\mathbf{V} \times \int_{-\infty}^{+\infty} \frac{(ku + \mu w)e^{ikx + i\mu z}}{[\lambda + i(ku + \mu w)](k^2 + \mu^2)}d\mu. \quad (8)$$

We will need to compute the corresponding potential perturbation

$$\delta\phi_1 = 8\pi ia\bar{m}G^2e^{\lambda t}\int\frac{F'(V)}{V}d\mathbf{V} \times \int_{-\infty}^{+\infty} \frac{(ku + \mu w)e^{ikx + i\mu z}}{[\lambda + i(ku + \mu w)](k^2 + \mu^2)}d\mu \quad (9)$$

only at $z = 0$. We use the spherical coordinates $u = V\sin\psi\sin\theta$, $v = V\cos\theta$, $w = V\cos\psi\sin\theta$ to integrate the right-hand side of (9) over velocity space. We first integrate over ψ. The integral reduces to a tabulated form via a change of the origin of ψ. We have

$$\int_0 \frac{k\sin\psi + \mu\cos\psi}{\lambda + iV\sin\theta(k\sin\psi + \mu\cos\psi)}d\psi = -\frac{2\pi iV(k^2 + \mu^2)\sin\theta}{\sqrt{\lambda^2 + V^2(k^2 + \mu^2)\sin^2\theta}[\lambda + \sqrt{\lambda^2 + V^2(k^2 + \mu^2)\sin^2\theta}]}. \quad (10)$$

We then multiply the resulting expression (10) by $\sin\theta$ (which enters the Jacobian) and integrate over θ to obtain

$$-4\pi i\left(1 - \frac{\lambda}{V\sqrt{k^2 + \mu^2}}\arctan\frac{V\sqrt{k^2 + \mu^2}}{\lambda}\right). \quad (11)$$

Using these intermediate manipulations, we can write the remaining integral over $\mathbf{V}$ and μ on the right-hand side of (9):

$$\delta\phi_1 = 32\pi^2\bar{m}aG^2e^{\lambda t + ikx} \times \iint\frac{VF'(V)}{(k^2 + \mu^2)^2}\left(1 - \frac{\lambda}{V\sqrt{k^2 + \mu^2}}\arctan\frac{V\sqrt{k^2 + \mu^2}}{\lambda}\right)dVd\mu. \quad (12)$$

We then integrate the right-hand side of (12) over V by parts. The result is easy to integrate over μ, yielding

$$\delta\phi_1 = \frac{32\pi^3\bar{m}aG^2}{\lambda^2}e^{\lambda t + ikx} \times \int_0^\infty\left(\frac{V}{\sqrt{k^2V^2 + \lambda^2}} - \frac{1}{k}\right)V^2F(V)dV. \quad (13)$$

The derivation of Eq. (6) implies that our calculations apply to the case of instability. In principle, we could study perturbations that are unstable even without a halo, to derive small corrections to the increment. However, of greater interest is the case where a perturbation in the absence of a halo has the form of a periodic wave, and the influence of the halo amplifies it. To analyze such slow pumping, we assume that $\lambda = \varepsilon - i\omega$,

where ω is real and ε is a small positive correction factor. Due to the way in which it was derived, the right-hand side of (13) is an analytical function of λ, so that we should chose the branch of the square root that is the analytical continuation from positive λ. We thus obtain

$$\left.\begin{aligned}\sqrt{k^2V^2+\lambda^2} &= \sqrt{k^2V^2-\omega^2}-\frac{i\varepsilon\omega}{\sqrt{k^2V^2-\omega^2}}+\ldots \quad (kV>\omega)\\ \sqrt{k^2V^2+\lambda^2} &= -i\sqrt{\omega^2-k^2V^2}+\frac{\varepsilon\omega}{\sqrt{\omega^2-k^2V^2}}+\ldots \quad (kV<\omega).\end{aligned}\right\} \tag{14}$$

The real part of the integral in (13), whose contribution is synchronous with the density variations, is of little interest and yields only small corrections to the pressure and self-gravitation in the gaseous disk. This becomes quite clear from Eqs. (18) below. We leave only the imaginary part, denoting its contribution to $\delta\phi_1$ as $\delta\tilde{\phi}_1$. In view of relations (14), we have

$$\delta\tilde{\phi}_1 = -\frac{32\pi^3 i\bar{m}aG^2}{\omega^2}e^{-i\omega t+ikx}\int_0^{\frac{\omega}{k}}\frac{V^3F(V)dV}{\sqrt{\omega^2-k^2V^2}} \tag{15}$$

(we ignore ε where it is negligible). Thus far, we have written expression (15) in a reference frame in which the halo is at rest. However, this restriction can easily be removed. Most importantly, the wave vector can, in principle, have an arbitrary direction. This generalization is easy to realize by replacing k in the integral (15) by $\sqrt{k_1^2+k_2^2}$, and kx in the exponent by the scalar product k_1x+k_2y, where k_1 and k_2 are the x and y components of the wave number vector, respectively. We further assume that the halo itself moves at a velocity (w_1, w_2) with respect to the disk. If x and y are the coordinates in the old reference frame (where the halo is at rest) then the coordinates in the new reference frame (where the halo moves) are $x' = x + w_1t$, $y' = y + w_2t$, $t' = t$. Correspondingly, the density perturbation (1) is expressed differently in the two frames:

$$\delta\sigma = ae^{-i\omega+ik_1x+ik_2y} = ae^{-i\omega t+ik_1(x'-w_1t)+ik_2(y'-w_2t)}.$$

Thus, we have a different frequency in the new reference frame:

$$\omega' = \omega + k_1w_1 + k_2w_2. \tag{16}$$

However, it follows from the derivation of (15) that this equation must contain the frequency in the old frame expressed in terms of the modified frequency ω'. As a result, we obtain from (15)

$$\begin{aligned}\delta\tilde{\phi}_1 &= -\frac{32\pi^3 i\bar{m}aG^2e^{-i\omega't+i(k_1x'+k_2y')}}{(\omega'-k_1w_1-k_2w_2)^2}\\ &\times\int_0^{V^*}\frac{V^3F(V)dV}{\sqrt{(\omega'-k_1w_1-k_2w_2)^2-k^2V^2}},\end{aligned} \tag{17}$$

where

$$V^* = \frac{|\omega'-k_1w_1-k_2w_2|}{k}.$$

We now derive linearized gasodynamic equations in the usual way, assuming that the gaseous disk rotates rigidly with angular velocity Ω. Introducing the sound speed c and velocity perturbations v_x and v_y, where the x- and y-axes are directed along the galactic radius and transversal, respectively, we find

$$\left.\begin{aligned}-i\omega v_x &= -\frac{c^2}{\sigma}ik_1\delta\sigma+\frac{2\pi Gik_1\delta\sigma}{\sqrt{k_1^2+k_2^2}}+2\Omega v_y-ik_1\delta\phi_1\\ -i\omega v_y &= -\frac{c^2}{\sigma}ik_2\delta\sigma+\frac{2\pi Gik_2\delta\sigma}{\sqrt{k_1^2+k_2^2}}-2\Omega v_x-ik_2\delta\phi_1\\ -i\omega\delta\sigma+\sigma(ik_1v_x+ik_2v_y) &= 0.\end{aligned}\right\} \tag{18}$$

Without the effect of the halo, this relation becomes the usual dispersion equation:

$$\omega^2 = 4\Omega^2 + c^2k^2 - 2\pi G\sigma k. \tag{19}$$

After introducing the abbreviation $Hi\delta\sigma$ ($H<0$) for the halo correction in (17), the dispersion equation (19) can be easily rewritten

$$\omega^2 = \omega_0^2 + k^2H\sigma i, \tag{20}$$

where $\omega_0^2 = 4\Omega^2 + c^2k^2 - 2\pi G\sigma k$.

According to (20), in the case of rotationally symmetric perturbations, the time dependence of the amplitude is described by the factor $\exp(-i\omega_0t + (k^2H\sigma/2\omega_0)t)$, so that the perturbations are always damped. The situation is quite different in the case of nonsymmetrical oscillations. Let us direct the local x and y axes along the radius and transversal of the galaxy, respectively. We then have $k_2 = m/R$, where R is galactocentric distance and m is an integer. Without loss of generality, we can assume that $k_1 > 0$, in which case $k_2 > 0$ or $m > 0$ and $k_2 < 0$ or $m < 0$ correspond to trailing and leading waves, respectively. Note that the frequency ω_0 in (20) is measured with respect to a rotating frame fixed to the gas. It is equal to the frequency ω' in Eq. (16).

For an observer fixed to the halo, the frequency is equal to the combination $\omega' - k_1 w_1 - k_2 w_2$, which has a different sign: it is negative. In this case, the requirements leading to formulas (14) are no longer valid: the sign of the second term on the right-hand side of (14) and, by continuity, the sign of the first, main term of (14) are reversed compared to the case above, thereby also implying that $H > 0$. Broadly speaking, galaxies exhibit negligible systematic radial motions, $w_1 = 0$.[1] On the other hand, $w_2 < 0$ (the rotation of the halo trails that of the disk). It follows that the criterion for reversal of the sign of H, and, consequently, the criterion for instability, is given by the inequality

$$\omega_0 - \frac{m}{R} w_2 < 0. \tag{21}$$

This criterion is satisfied when $m < 0$, i.e., for leading spirals with sufficiently large $|m|$.

To obtain realistic calculations, we give below several examples of calculations of the integral in (15). We have for a Gaussian distribution $F(V) = \alpha \exp(-sV^2)$ with arbitrary parameters α and s

$$\int_0^\varphi \frac{V^3 F(V)}{\varphi^2 - V^2} dV = \alpha \left[\left(\varphi^2 + \frac{1}{2s} \right) e^{-s\varphi^2} \int_0^\varphi e^{sy^2} dy - \frac{\varphi}{2s} \right],$$

$$(\varphi = V\sqrt{k_1^2 + k_2^2}).$$

Here is another example. If

$$F(V) = \frac{\beta}{(b^2 + V^2)^2},$$

then

$$\int_0^\varphi \frac{V^3 F(V)}{\sqrt{\varphi^2 - V^2}} dV = \frac{\beta}{4(b^2 + \varphi^2)^{3/2}}$$

$$\times \left[-2\varphi\sqrt{b^2 + \varphi^2} + (b^2 + 2\varphi^2) \ln \frac{\sqrt{b^2 + \varphi^2} + \varphi}{\sqrt{b^2 + \varphi^2} - \varphi} \right].$$

Concerning the transformations finally leading us to Eq. (15), note that the multiple integration can be performed in a different sequence. For example, the result is the same if we begin by integrating over μ and then proceed to integrating over the velocity components.

3. A STELLAR DISK

We now address the case of a stellar disk in much the same way as we analyzed that of a gaseous disk above. Consider a stellar system with a flat Maxwellian velocity distribution h^2 in each direction and with surface mass density γ:

$$f_0 = \frac{\gamma}{2\pi R^2 \bar{m}} e^{-\frac{v_x^2 + v_y^2}{2h^2}}. \tag{22}$$

The perturbed phase-space density f evolves in accordance with the Boltzmann equation with allowance for rotation [6]:

$$\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + \left(-\frac{\partial \phi}{\partial x} + \Omega^2 x + 2\Omega v_y \right) \frac{\partial f}{\partial v_x} + \left(-\frac{\partial \phi}{\partial y} + \Omega^2 y - 2\Omega v_x \right) \frac{\partial f}{\partial v_y} = 0. \tag{23}$$

Linearizing Eq. (23) yields

$$\frac{\partial \delta f}{\partial t} + v_x \frac{\partial \delta f}{\partial x} + 2\Omega v_y \frac{\partial \delta f}{\partial v_x} - 2\Omega v_x \frac{\partial \delta f}{\partial v_y} - \left(\frac{\partial \delta\phi}{\partial x} + \frac{\partial \delta\bar{\phi}_1}{\partial x} \right) \frac{\partial f_0}{\partial v_x} = 0. \tag{24}$$

Here, we again assume that the wave propagates in the x direction with increment λ and wave number k. After separating the variables x and t, we reduce equation (24) to the form

$$(\lambda + ikv_x)\delta f + 2\Omega \left(v_y \frac{\partial \delta f}{\partial x} - v_x \frac{\partial \delta f}{\partial y} \right) = ik(\delta\phi + \delta\bar{\phi}_1) \frac{\partial f_0}{\partial v_x}. \tag{25}$$

On the whole, Eq. (25) is identical to the well-known equation of density-wave theory [7], and differs only in the additional term $\delta\bar{\phi}_1$. We solve this equation using the usual techniques, simplifying them somewhat in the process. We use the Fourier transform in the same way as in the gaseous-disk case:

$$\xi(\mu_x, \mu_y) = \iint \delta f e^{i(\mu_x v_x + \mu_y v_y)} dv_x dv_y. \tag{26}$$

As a result, we obtain from (25)

$$\lambda\xi + k \frac{\partial \xi}{\partial \mu_x} + 2\Omega \left(\mu_y \frac{\partial \xi}{\partial \mu_x} - \mu_x \frac{\partial \xi}{\partial \mu_y} \right) = \frac{k\gamma\mu_x}{\bar{m}} (\delta\phi + \delta\bar{\phi}_1) e^{-\frac{h^2}{2}(\mu_x^2 + \mu_y^2)}. \tag{27}$$

We now introduce polar coordinates in the space of the auxiliary variables μ_x and μ_y:

$$\mu_x = -s\sin\theta, \quad \mu_y = s\cos\theta - \frac{k}{2\Omega}. \tag{28}$$

[1] Although some cases of intense radial motion are known, they are beyond the domain of applicability of density-wave theory.

Eq. (27) can then be written

$$\lambda\xi - 2\Omega\frac{\partial\xi}{\partial\theta} = -\frac{\gamma s\sin\theta}{\bar{m}}(\delta\phi + \delta\bar{\phi}_1)e^{-\frac{h^2}{2}\left(s^2 - \frac{ks\cos\theta}{\Omega} + \frac{k^2}{4\Omega^2}\right)}. \quad (29)$$

This is a linear differential equation in standard form. We multiply it by $\exp(-\lambda\theta/2\Omega)$ and integrate both the left-hand and right-hand sides from zero to an arbitrary θ:

$$2\Omega\xi e^{-\frac{\lambda\theta}{2\Omega}} = \frac{\gamma s(\delta\phi + \delta\bar{\phi}_1)}{\bar{m}}e^{-\frac{h^2}{2}\left(s^2 + \frac{k^2}{4\Omega^2}\right)} \times\left[\int_0^\theta \sin\theta e^{-\frac{\lambda\theta}{2\Omega} + \frac{h^2ks\cos\theta}{2\Omega}}d\theta - c_1\right], \quad (30)$$

where the integration constant c_1 is determined from the condition that changing $\theta = 0$ to $\theta = 2\pi$ on the left-hand side of (30) should be equivalent to multiplying by the constant factor $\exp(-\pi\lambda/\theta)$. We thus have

$$c_1 = \frac{\int_0^{2\pi}\sin\theta e^{-\frac{\lambda\theta}{2\Omega} + \frac{h^2ks\cos\theta}{2\Omega}}d\theta}{1 - e^{-\frac{\pi\lambda}{\Omega}}}. \quad (31)$$

We are mainly interested in the stellar density increment, which can be obtained from (26) using the substitutions $\mu_x = \mu_y = 0$ or $\theta = 0$, $s = k/2\Omega$. According to Eq. (30), we have in this case

$$\bar{m}\iint\delta f dv_x dv_y = -\frac{\gamma k}{4\Omega^2}(\delta\phi + \delta\bar{\phi}_1)e^{-\frac{h^2k^2}{4\Omega^2}}c_1. \quad (32)$$

If the increment $\lambda = -i\omega$ is purely imaginary, expression (31) can be transformed via a simple substitution into the form

$$c_1 = \frac{\int_0^{\pi}\sin\lambda\,\sin\frac{\omega\lambda}{2\Omega}e^{-\frac{h^2k^2\cos\lambda}{4\Omega^2}}d\lambda}{\sin\frac{\pi\omega}{2\Omega}}. \quad (33)$$

We introduce $\nu = \omega/2\Omega$ to write the equation in a shorter form. In virtually all the cases considered, the increment $\nu < 1$ and we adopt this assumption throughout all further analysis. It then follows from definition (33) that $c_1 > 0$. Note that c_1 increases monotonically with ω, provided all other parameters are constant. To prove this, it is sufficient to prove that the function

$$\varphi(\nu) = \frac{\sin\lambda\nu}{\sin\pi\nu}$$

increases monotonically for all allowed $\lambda \in (0, \pi)$. Indeed,

$$\frac{\varphi'(\nu)}{\varphi(\nu)} = \lambda\cot\lambda\nu - \pi\cot\pi\nu, \quad (34)$$

and for any allowed λ, differentiating (34) yields

$$\frac{\partial}{\partial\lambda}\frac{\varphi'(\nu)}{\varphi(\nu)} = \cot\lambda\nu - \frac{\nu\lambda}{\sin^2\nu\lambda} = \frac{\sin 2\lambda\nu - 2\lambda\nu}{2\sin^2\nu\lambda} < 0.$$

Since the right-hand side of (34) vanishes at the extreme value $\lambda = \pi$, it is positive throughout the entire remaining interval $0 < \lambda < \pi$, which was what we wished to prove.

Without a halo, formulas (32) and (33) combined with the relation between the surface density and potential perturbations yield the dispersion relation

$$\frac{\pi G\gamma}{2\Omega^2}c_1 = e^{\frac{h^2k^2}{4\Omega^2}}, \quad (35)$$

which obviously coincides with the well-known result [7–11] for the case of rigid rotation, up to a simple integration by parts.

The correction term $\delta\bar{\phi}_1$ is included in the same way as in the gaseous disk case. The key point here is the aforementioned monotonic dependence of the left-hand side of Eq. (35) on ω, which plays the same role in determining the correction to λ as it does in the simpler dispersion Eq. (20). The specific formulas for this correction are more cumbersome; however, the qualitative conclusions about stability or instability remain unchanged.

4. VERTICAL OSCILLATIONS

Thus far, we have analyzed exclusively oscillations in the plane of the disk itself. However, dynamical friction has meaning for vertical, bending oscillations as well. Here, we restrict our analysis to the case of a gaseous disk, since membrane instability makes the problem more complicated for stellar disks [7].

As above, we consider vertical displacement in the form

$$\eta = ae^{\lambda t + ikx}. \quad (36)$$

In a linear approximation, such bending is similar to the formation of a double layer with density $\eta\sigma$, or, in three-dimensional terms, an additional density appears:

$$\delta\rho = -\eta\sigma\bar{\delta}'(z) = \frac{i\mu}{2\pi}ae^{\lambda t}\int_{-\infty}^{+\infty}e^{ikx + i\mu z}d\mu. \quad (37)$$

The density perturbation (37) corresponds to the potential perturbation

$$\delta\phi = 2iaGe^{\lambda t}\int_{-\infty}^{+\infty}\mu\frac{e^{ikx+i\mu z}}{k^2+\mu^2}d\mu. \quad (38)$$

As above, the increment of the phase-space density is given by the Boltzmann Eq. (5), and we obtain the following additional potential due to the halo:

$$\delta\phi_1 = -8\pi a\bar{m}G^2e^{\lambda t}\int\frac{F'(V)}{V}d\mathbf{V} \times\int_{-\infty}^{+\infty}\frac{\mu(ku+\mu w)e^{ikx+i\mu z}}{[\lambda+i(ku+\mu w)](k^2+\mu^2)^2}d\mu. \quad (39)$$

Compared to the case of horizontal oscillations, we must first determine the additional component of the vertical acceleration by differentiating (39) with respect to z, then substituting $z = 0$. Integration of (39) over all variables in the same sequence as above yields the additional vertical acceleration

$$W_3 = -\frac{32\pi^3a\bar{m}G^2e^{\lambda t+ikx}}{\lambda^2}\int_0^{+\infty}(\sqrt{\lambda^2+k^2V^2}-kV)VF(V)dV, \quad (40)$$

to be added to the main vertical acceleration components W_1 and W_2, due to bending of the layer itself and the unperturbed halo field, respectively. The force W_1 is given by the well-known formula [12] $W_1 = -2\pi Gk\sigma\eta$, and we adopt $W_2 = -2\pi G\rho^*\eta$, where ρ^* is some effective halo density. Leaving only the imaginary part in Eq. (40), we obtain the equation of motion

$$\frac{\partial^2\eta}{\partial t^2} = -\omega_0^2\eta + \frac{32\pi^3\bar{m}G^2\eta i}{\omega_0^2} \times\int_0^{\frac{\omega_0}{k}}\sqrt{\omega_0^2-k^2V^2}VF(V)dV. \quad (41)$$

The unperturbed frequency ω_0 is given by $\omega_0^2 = 2\pi G(\rho + k\sigma)$. Equation (41) can be generalized to the case of arbitrary wave orientation in the same way as was done for horizontal oscillations, and criterion (21) remains valid in this case.

5. CONCLUSIONS

The quantity H in (20), computed using (15), can be estimated more or less accurately for the case of our Galaxy. Let us now substitute, for example, the values adopted in [7]: $k = 2\pi/\lambda$, $\lambda = 3.5$ kpc, $\sigma = 9m_\odot\text{pc}^{-2}$, and $\omega' - k_2w_2 = 2\Omega_p = 24$ km s^{-1} kpc^{-1} (Ω_p is the rotation speed of the spiral pattern). The rms velocity in the halo (which is, strictly speaking, referred to as a thick disk in [13]) depends only slightly on the direction, and we, therefore, take it to be equal to 104 km/s, on average. Buser and Rong [14] give a stellar density of 0.054 pc^{-3}. We assume the mean mass of halo stars to be $0.4m_\odot$ [15, 16]. We now substitute the Maxwellian exponent into (15) and assume it to be equal to its central value, since V^* is sufficiently small to justify this approximation. Simple calculations yield $k^2H\sigma \sim 1\times10^{-33}$ s^{-2}, which is three orders of magnitude smaller than the square of the angular velocity of the Galaxy, Ω^2. However the halo correction in (20) should be compared not with Ω^2, but with ω_0^2, a quantity that could be substantially smaller if the Galaxy is close to the instability limit (there is currently no consensus about how close it actually is). The very rough estimate $\omega_0^2 = 0.1\Omega^2$ yields an imaginary correction to frequency ω with a magnitude of 0.005, which corresponds to a relative increase of ~0.03 over one period. However, it takes about ten periods for the wave to propagate from the center to the periphery, so that the amplitude changes substantially (by ~30%) over a single cycle.

Note that, to some approximation (i.e., if ω' = const), the correction depends on k as k^{-2}, so that open spirals should be more sensitive to the effect of a halo. Accordingly, theoretical mechanisms for the amplification and damping of perturbations leading to such open spirals should be refined.

Similar effects seem to appear in numerical simulations, provided they involve the formation of more or less well-defined relatively high-velocity systems. The results of such simulations should be interpreted with caution.

A similar amplification of vertical oscillations involving a superposition of several modes and nonlinear effects could increase the disk thickness until turbulence and magnetic viscosity compensate the instability. On the whole, allowance for dynamical friction will probably change our view of the evolution of galactic disks substantially. In the future, we plan to analyze other, resonant, effects, which differ from the dynamical friction considered above in their high degree of regularity [4]. However, such problems require a global rather than local approach. We note, by the way, that the criterion (21) derived in this paper is very similar to the well-known Landau condition for inverse damping [17, 18].

ACKNOWLEDGMENTS

We are sincerely grateful to A.M. Fridman for his careful review and valuable comments.

REFERENCES

1. S. Chandrasekhar, Astrophys. J. **97**, 255 (1943); **98**, 54 (1943).
2. A. S. Baranov and Yu. V. Batrakov, Astron. Zh. **51**, 310 (1974) [Sov. Astron. **18**, 180 (1974)].

3. A. S. Baranov, Astron. Zh. **61**, 1098 (1984) [Sov. Astron. **28**, 642 (1984)].
4. A. S. Baranov, Astron. Zh. **68**, 1160 (1991) [Sov. Astron. **35**, 582 (1991)].
5. V. A. Antonov, E. I. Timoshkova, and K. V. Kholshevnikov, *Introduction to the Theory of Newtonian Potentials* (Nauka, Moscow, 1988).
6. C. Pichon and R. C. Cannon, Mon. Not. R. Astron. Soc. **291**, 616 (1997).
7. V. L. Polyachenko and A. M. Fridman, *Equilibrium and Stability in Gravitating Systems* (Nauka, Moscow, 1976).
8. A. J. Kalnajs, Astrophys. Space Sci. **13**, 279 (1971).
9. C. C. Lin and S. H. Frank, Proc. Natl. Acad. Sci. **55**, 229 (1966).
10. R. Wielen, Mitt. Astron. Ges., No. 30, 31 (1971).
11. A. S. Marochnik and A. A. Suchkov, Astron. Zh. **46**, 319 (1969) [Sov. Astron. **13**, 252 (1969)].
12. C. Hunter and A. Toomre, Astrophys. J. **155**, 747 (1969).
13. S. Bartašiūte, Balt. Astron. **3**, 16 (1994).
14. R. Buser and J. Rong, Balt. Astron. **4**, 1 (1995).
15. S. M. Faber and H. B. French, Astrophys. J. **235**, 405 (1980).
16. J. H. Lacy, C. H. Townes, and D. J. Hollenbach, Astrophys. J. **262**, 120 (1982).
17. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
18. A. M. Fridman and V. L. Polyachenko, *Physics of Gravitating Systems* (Springer, New York, 1984), Vols. I and II.

Translated by A. Dambis

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 654–658. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 737–742.

Studies of Weak Compact Radio Sources

S. A. Tyul'bashev and O. I. Malov

Pushchino Radio Astronomy Observatory, Astro Space Center, Lebedev Institute of Physics, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia

Received October 20, 1999

Abstract—Approximately 20% of weak sources in a scintillation survey at 102 MHz were not identified in other, more sensitive, low-frequency surveys. These sources had very high compactness and steep spectra. Since, as shown earlier, these sources are probably quasars, the epoch of the radio birth of quasars should correspond to flux densities of ~0.01 Jy. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

During a 102-MHz scintillation survey [1] of the first area of the 7C survey [2], 125 scintillating sources were detected. In the vast majority of cases, sources with 102-MHz flux densities greater than 0.5 Jy were identified with sources in other low-frequency radio catalogs [3], or else there were several candidates and additional studies were required to ensure an unambiguous identification. However, the situation was substantially worse for sources with flux densities lower than 0.4 Jy. An appreciable fraction of these sources were not identified; some of these sources have high compactness ($R \simeq 1$) and steep spectra ($\alpha > 1 S \sim \nu^{-\alpha}$). The 102-MHz scintillating-source survey was complete to 0.4 Jy, and the flux densities of the weakest sources were 0.15 Jy. The 7C survey at 151 MHz, which we adopted as a basis for our own survey, is complete to about 0.2 Jy, and the weakest sources in the catalog have flux densities of approximately 0.06–0.08 Jy [2]. This means that the sensitivity of the 7C survey was nearly twice that of our own. The mean compactness of the identified radio sources was 0.4 [3], so that a source with a flux density of 0.15 Jy in a compact (<1″) component should, on average, have an integrated flux density of about 0.35–0.4 Jy, and should be easy to identify with radio sources in sensitive surveys.

There are a number of possible explanations for the fact that some of our compact radio sources could not be identified with radio sources in other surveys with sensitivities that are either comparable to or exceed that of our own (after translating the flux density from the observed frequency to 102 MHz). In this paper, we consider several of these possibilities.

2. A PULSAR NATURE FOR THE WEAK SOURCES

It was shown in [4] that there are no known pulsars in the first area of the 7C survey, but there could be one or several new pulsars. This area was surveyed for pulsars only once, during a search for long-period ($P > 0.2^s$) pulsars [5]. The sensitivity of this survey was 10 mJy at 400 MHz. Note that the sensitivity of a pulsar survey depends on dispersion measure and the relationship between the characteristic pulse duration and the period of the pulsar. In scintillation surveys, pulsars are observed to be compact sources with continuous emission. Consequently, a pulsar with an integrated flux density of 0.2 Jy at 102 MHz should be detected as a scintillating source, even if its period is 1 ms and its pulse occupies a large fraction of the period. The sensitivities of the pulsar survey [5] and our scintillation survey were comparable. For example, a pulsar with a flux density of 10 mJy at 400 MHz should have a 100-MHz flux density of 0.1 Jy for $\alpha = 1.5$ and 0.3 Jy for $\alpha = 2.5$. As mentioned in the Introduction, the typical flux densities for the unidentified scintillating sources were less than 0.4 Jy. Therefore, we undertook a pulsar search in order to directly estimate their possible contribution to the scintillating-source survey.

The table gives a list of sources investigated as pulsar candidates. From the scintillating-source catalog [1], we selected sources that were unidentified in other catalogs or had pulsar-like spectra. The first column of the table gives the name of the source from [1], the second the 102-MHz flux density, and the third an estimate of the spectral index between 151 and 102 MHz, assuming that the compactness of the source is equal to unity. If there was no identification, we give a lower limit for α based on the assumed value $S_{151} = 0.2$ Jy. The fourth column contains comments. The letter A indicates that the source was at the edge of the studied field, and therefore its coordinates and flux density may be uncertain. The letters B and C indicate that the lower limit on α between 151 MHz (the 7C catalog [2]) and 365 MHz/408 MHz (Texas data [6]/Bologna data [7]) is greater than one or two, respectively. We should point out individually the source 0604 + 4956, which has a flux density of 0.47 Jy (7C) and is absent from other catalogs ($\alpha_{151-365} \simeq 2$). The fifth and sixth columns of the table indicate how many days the source was observed in "second" and "millisecond" regimes, respectively.

1063-7729/00/4410-0654$20.00

Pulsar candidates

Source name	S_{102}, Jy	Estimate of α	Comments	Observations in the "second regime," days	Observations in the "millisecond regime," days
1	2	3	4	5	6
0545 + 4423	0.26	≥0.7	A	6	3
0546 + 4246	0.39	≥1.7	A	4	3
0546 + 4222	0.26	≥0.7	A	4	3
0553 + 4336	0.37	1.3		5	3
0555 + 4503	0.27	~0	B	5	3
0555 + 4849	0.2	~0		7	3
0602 + 4526	0.56	≥2.6		7	3
0604 + 4850	0.19	~0		3	3
0604 + 4116	0.18	~0		3	2
0604 + 4956	0.2	–2.2		6	2
0606 + 3940	0.8	≥3.5	A	1	1
0608 + 3954	0.59	1.2		5	3
0609 + 4327	0.41	1.7		5	3
0610 + 4510	0.53	–0.29	C	4	3
0612 + 4851	0.2	~0		4	3
0614 + 5006	0.27	0.9		9	3
0622 + 4612	0.19	~0		9	3
0638 + 4839	0.19	~0		6	3
0639 + 4506	0.13	~0		7	3
0643 + 4346	0.19	~0		10	3
0648 + 4024	0.26	≥0.7		9	3
0653 + 4038	0.47	1.6		10	4
0657 + 4441	0.33	≥1.3		6	4
0658 + 4941	0.3	~0		6	2
0704 + 3950	0.5	1.4		10	3
0708 + 4828	0.27	–1.4	B	9	5

The observations were carried out using two receivers designed for pulsar observations. The first (AC-32) had a channel bandwidth of 5 kHz. Due to the presence of strong interference on the Large Phased Array, we were able to use no more than 16 channels; i.e., our bandwidth was <100 kHz, which is narrow for searches. Our observations using this receiver were obtained from November 1996 through March 1997. A last series of observations was carried out in May 1999 (after which Large-Phased-Array operation was moved to 111 MHz) using a wide-band spectral analyzer with a 20-kHz channel bandwidth (AC-128). We recorded 32 channels, so that the total bandwidth was 640 kHz. The observations were carried out in various regimes: on the AC-32 receiver with time constants 0.3072, 2.56, and 5.12 ms (using 16 channels) and on the AC-128 receiver with time constants 5.12 and 20.48 ms (using 16 and 32 channels).

The basis of the search program was the construction and visual monitoring of the power spectrum, performing an automated search for peaks in the power spectrum exceeding the 7σ level. During the search regime, we also monitored the recordings for known pulsars. Figure 1a shows the power spectrum of the strong pulsar 0823 + 26, in which a large number of harmonics are visible. The pulsar period derived from the power spectrum is P = 0.52979 s, while the true period is P_{true} = 0.53066 s. Thus, for strong pulsars, the error in the derived period does not exceed 0.2%. Figure 1b presents the power spectrum of the weak pulsar 0919 + 06: only the main harmonic of the pulsar reaches the 7σ level. The pulsar period estimated from this harmonic is P = 0.4207 s, while P_{true} = 0.4306 s. Thus, the error in the derived periods of weak pulsars is better than 2.5%. Since the integrated flux density of a weak pulsar is comparable to that of the weakest of the

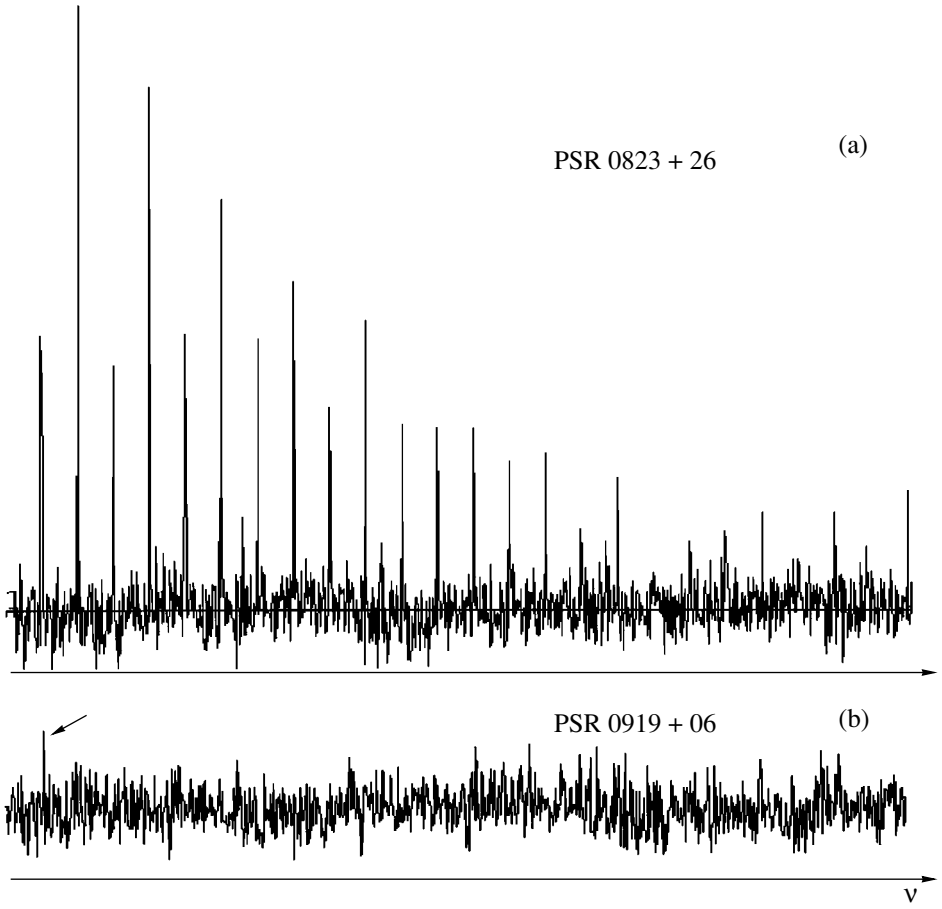


Fig. 1. (a) Power spectrum for the pulsar 0823 + 26, in which about 20 harmonics are visible. The integrated flux density of the pulsar is S_{102} = 620 mJy [8]. (b) Power spectrum for the pulsar 0919 + 06. The arrow marks a single harmonic detected at the 7σ level. S_{102} = 128 mJy [8].

observed sources, we expected that spectral lines corresponding to pulsar periods should easily be detected in the power spectra. Not a single pulsar was detected. Consequently, we conclude that the contribution of pulsars with periods >10 ms to the scintillating-source survey is negligibly small.

3. A NEW POPULATION OR "OLD FRIENDS"?

Thus, we are dealing with a population of very compact (scintillating), weak radio sources with steep spectra. A source with a compact flux density of 0.25 Jy (the typical flux density for the compact components of the unidentified sources in our survey) should have, on average, an integrated flux density of 0.6 Jy (for a typical compactness of 0.4 [3]). Consequently, since the 7C survey is complete to 0.2 Jy [2], a lower limit to the spectral index of such a source is $\alpha \geq 3.5$. On the other hand, if the compactness of the weak ($S_{102} \sim 0.2$ Jy), unidentified sources is unity, their spectral indices will be about 0.7. The earlier study [3] indicates that the integrated spectra for identified scintillating sources are steep: $\bar{\alpha} = 0.84$. It was also shown that the spectra of the compact components of a large fraction (60%) of the identified radio sources were steep: $\bar{\alpha} > 0.5$. We also noted that, on average, weaker identified sources are more compact.

Figure 2 shows the dependence of the source compactness on integrated flux density, including both compactness estimates for identified scintillating sources and lower limits to the compactness of unidentified sources from [3]. In total, the figure includes 102 compactness estimates. The growth in the mean compactness with decreasing flux density in the compact component can clearly be seen. There is a large scatter in the compactness estimates about their mean values at given flux densities. This is probably due to the fact that active galactic nuclei were born at various cosmological epochs. This means that active nuclei at different stages of their evolution, i.e., having different compactness, can have the same flux density.

It is interesting that the central points in the distribution lie on a fairly linear trend, suggesting that the compactness grows smoothly as the compact flux density decreases. An extrapolation of this trend suggests that

the vast majority of scintillating sources with very low flux densities ($S_{102} \sim 0.01$ Jy) will have $R = 1$. Such sources should be very distant, corresponding to the initial epoch of the birth of active galactic nuclei at radio wavelengths. Based on the formula for the photometric distance to an object and known redshifts, we can make a very rough estimate of the distances to such sources: $z \sim 20 \pm 10$ (for $q_0 = 1/2$). It is most likely that the weak scintillating sources are either young radio galaxies that have not yet expanded beyond the confines of their host galaxy (so that their angular sizes are small) or quasars. However, investigation of the compactness of various classes of objects (quasars, giant radio galaxies, Seyfert galaxies, BL Lacertae objects, etc.) has shown that the most likely possibility is that these sources are quasars [3]. The mean compactness of the identified scintillating sources and its rms deviation coincide very well with those for quasars observed in earlier scintillation studies. It appears that the optical identifications for weak ($S \sim 0.01$ Jy) radio sources with small angular size ($\theta \sim 0.1''$) and high compactness ($R \sim 1$) at low frequencies could correspond to the epoch of the birth of the first radio quasars.

The introduction of a new class of sources whose nature is as yet unknown could provide an alternative explanation for the weak, unidentified scintillating sources. These sources could have very high compactness ($R \sim 1$), very steep integrated spectra ($\alpha > 1.5$–2), or intermediate characteristics. The possibility of the emergence of a new class of radio sources is not ruled out. Recall, for example, the discovery of the class of compact steep-spectrum sources in 1982 [9, 10]. However, this explanation seems much less plausible to us than that discussed earlier in this Section.

4. INCOMPLETENESS OF THE 7C SURVEY

According to [2], the 7C survey is complete to 0.2 Jy. Consequently, sources from our survey, which is complete to 0.4 Jy, should be easily identified in the 7C catalog. Therefore, the lack of identifications for some sources could suggest that the 7C survey is, in reality, not complete to the claimed flux-density level of 0.2 Jy. Indeed, our surveys of the first and second areas of the 7C survey have detected a number of scintillating sources that clearly fall in these areas, were not identified in the 7C catalog, but were identified in the less sensitive 6C catalog [11]. There were eight such sources in these two areas. The most striking example is the source 1045 + 3519 [12], whose scintillating component has a flux density of 2 Jy at 102 MHz and which had an integrated flux density of 3 Jy at 151 MHz in the 6C catalog [11], but is absent from the 7C catalog. Since fewer than >10% of the sources in the 7C survey scintillate, it is easy to estimate that at least ~2% of sources in the 7C survey with flux densities $S_{151} \geq 0.2$ Jy could be lost.

In our opinion, this represents a fully admissible divergence between model computations and real observations. However, if all the unidentified sources were lost from the 7C survey due to its incompleteness, this suggests that 10–15% of sources with $S_{151} \geq 0.2$ Jy are missing from the 7C survey. Such substantial losses would lead to an appreciable distortion of the source counts [2] at flux densities 0.2–0.6 Jy, which is not observed. Consequently, incompleteness of the 7C survey cannot explain the large number of unidentified weak scintillating sources.

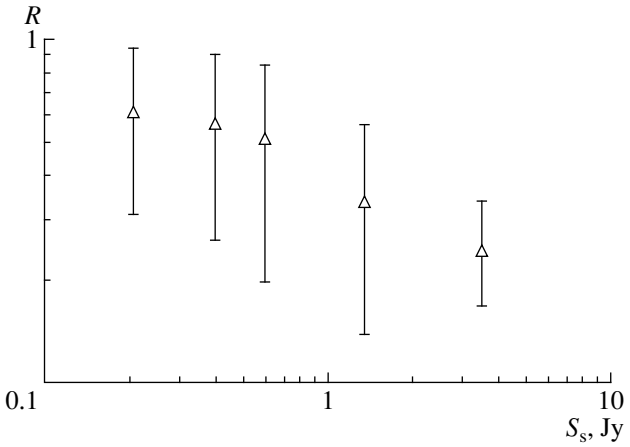


Fig. 2. Dependence of the mean compactness of scintillating sources on their flux density.

5. CONFUSION

The effect of confusion for the scintillating sources could also explain the appearance of unidentified sources. Indeed, one weak scintillating source could be observed in place of two or more weak and relatively closely spaced sources in the 7C survey area. This situation could easily arise, since it turned out that a substantial fraction ($\simeq$20%) of the scintillating sources had ambiguous identifications. Some of the scintillating sources in the first 7C survey area have up to four 7C candidate identifications. This effect was considered in the interpretation of the results of the scintillating-source survey [4]. It was shown that its role was small. Of course, we cannot rule out the possibility that it is acting in a few isolated cases.

Here, it is appropriate to say a few words about the possibility of incorrect determination of the coordinates of the unidentified sources. The identifications were carried out over a $\pm 15'$ area around the derived coordinates for the scintillating source. If the error in these coordinates were larger than this, the source would remain unidentified. However, this is improbable for two reasons. First, the coordinates for the vast majority of the identified sources differ from the 7C coordinates by no more than $5'$–$7'$, and there is no basis to believe that this accuracy should be appreciably worse for the unidentified sources. Second, in order to be able to be sure of our results and exclude such uncer-

tainties, we observed the same declination 3–5 times, giving us confidence in the resulting coordinates.

6. CONCLUSION

We have considered a number of possible reasons for the detection of a large number of scintillating sources (≈20% of all sources in our survey) that remain unidentified with sources in more sensitive surveys. The contributions of pulsars and confusion to the scintillating-source survey are negligibly small, or even entirely absent. Inaccurate determination of the coordinates for the scintillating sources hindering their identification is also improbable. The 7C survey, which we adopted as the basis for our observations, is undoubtedly partially incomplete, but it is unlikely that this can explain an appreciable number of the unidentified sources.

The emergence of a new population of sources with very high compactness ($R \sim 1$) and steep spectra ($\alpha \sim 1$) cannot be ruled out. Additional studies are required, especially optical identifications of these radio sources. However, we consider the existence of a new population of sources that is not manifest in other sensitive low-frequency surveys to be improbable.

Thus, the most probable explanation for the unidentified scintillating sources is that these weak sources are distant quasars with steep ($\alpha \sim 1$) spectra and high compactness ($R \sim 1$).

ACKNOWLEDGMENTS

We would like to thank V.M. Malofeev for his constant attention to this work and for useful discussions, and V.D. Pugachev for consultations about the software for the pulsar receivers. The pulsar studies were supported by an INTAS grant (96-0154) and the Russian Foundation for Basic Research (project no. 97-02-17372).

REFERENCES

1. V. S. Artyukh and S. A. Tyul'bashev, Astron. Zh. **73**, 661 (1996) [Astron. Rep. **40**, 601 (1996)].
2. Mark M. McGilchrist, J. E. Baldwin, J. M. Riley, *et al.*, Mon. Not. R. Astron. Soc. **246**, 110 (1990).
3. S. A. Tyul'bashev, Astron. Zh. **74**, 812 (1997) [Astron. Rep. **41**, 723 (1997)].
4. V. S. Artyukh and S. A. Tyul'bashev, Astron. Zh. **75**, 655 (1998) [Astron. Rep. **42**, 576 (1998)].
5. M. Damashek, J. H. Taylor, and R. A. Hulse, Astrophys. J. Lett. **225**, L31 (1978).
6. J. N. Douglas, F. N. Bash, F. A. Bozyan, *et al.*, Astron. J. **111**, 1945 (1996).
7. A. Ficarra, G. Grueff, and G. Tomassetti, Astron. Astrophys., Suppl. Ser. **59**, 255 (1985).
8. V. M. Malofeev, O. I. Malov, and N. V. Shchegoleva, Astron. Zh. **77**, 499 (2000) [Astron. Rep. **44**, 436 (2000)].
9. V. K. Kapahi, Astron. Astrophys., Suppl. Ser. **43**, 381 (1981).
10. J. A. Peacock and J. V. Wall, Mon. Not. R. Astron. Soc. **198**, 843 (1982).
11. S. E. G. Hales, J. E. Baldwin, and P. J. Warner, Mon. Not. R. Astron. Soc. **263**, 25 (1993).
12. V. S. Artyukh, S. A. Tyul'bashev, and E. A. Isaev, Astron. Zh. **75**, 323 (1998) [Astron. Rep. **42**, 283 (1998)].

Translated by D. Gabuzda

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 659–664. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 743–749.

Turbulent Diffusion in the Interstellar Medium

V. G. Lamburt, D. D. Sokolov, and V. N. Tutubalin

Moscow State University, Moscow, Vorob'evy gory, 119899 Russia

Received October 10, 1999

Abstract—We have calculated the coefficient of turbulent diffusion in a random flow with time restoration, describing the interstellar medium. Such a flow abruptly loses its memory at random times, forming a Poisson flow of events. The coefficient of turbulent diffusion in the flow is determined by the rms velocity and correlation time, as in mixing-length theory, but the numerical coefficient differs from that predicted by this theory. The closure equation derived by us for the transport of the mean concentration of a passive scalar takes a more complicated form than obtained in standard mean-field theory, but the main properties of the equation retain their validity. The possibility of extending the results of this exactly solved problem to the problem of transport in the turbulent interstellar medium is discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The transport of a passive scalar (temperature, dust) or a vector (magnetic field) admixture in the interstellar medium is determined largely by turbulent diffusion. In the simplest case, the equation for the transport of the mean concentration is identical in form to the true equation for transport of the admixture, but the diffusion coefficient changes, becoming equal to the coefficient of turbulent diffusion $D_T = \frac{1}{3} l^* v^*$, where v^* and l^* are the rms velocity and correlation length of the flow, respectively. Assuming that the memory time t^* is determined by the turnover time l^*/v^* of a turbulent vortex, this formula can be rewritten $\beta = \frac{1}{3}(v^*)^2 t^*$. These simple concepts are usually called mixing-length theory.

The equation of transport of a mean quantity can be more complicated than the corresponding microscopic equation. For example, a new term related to the so-called α effect, or the mean spirality, appears in the equation of transport of the mean magnetic field. However, simple formulas based on the concepts of mixing-length theory (see, e.g., [1, 2]) can be proposed for the corresponding transport coefficient.

Of course, these simple concepts are very crude, and the degree to which they correspond to properties of the real turbulent interstellar medium has long been the subject of discussion. For example, it has been argued in some studies (see, for example, [3]) that the turbulent transport coefficients for an infinitesimal magnetic field and for a very small (compared to the case of uniform distribution) magnetic field are very different. It was also pointed out in [4] that, when the memory time differs greatly from the vortex turnover time, this can considerably change the formulas for the turbulent transport coefficients. Beck *et al.* [5] hold the more conservative view that the formulas of mixing-length theory are fairly stable, and that traditional mean-field theory provides a crude agreement with observations.

In this study, we consider the effect of time correlations in the random velocity field on the turbulent transport coefficients. Numerical experiments in this area present serious problems, because it is difficult to model a medium with such large Reynolds numbers as those for the turbulent interstellar medium (the current state of the art in this area is described in [6]). Investigations based on various closure hypotheses confirm the important role of time correlations for descriptions of turbulent diffusion [7, 8]. However, these studies are based essentially on hypotheses of correlation splitting, and some exactly solved examples [9, 10] show that subtle effects can exist, at least for magnetic fields, which qualitatively change the behavior of admixtures in the case of very long temporal (or spatial) correlations. As a consequence, the construction of models of a turbulent flow that solve for the transport of a passive admixture without invoking correlation-splitting hypotheses takes on special importance. Such an approach is made possible by the use of functional integrals, which enable one to write the integral representation of a solution in which, in some cases, correlations are split exactly rather than approximately [11].

Thus far, it has been possible in this way to study short-correlation models, in which the memory time is assumed infinitely short [12], and models with resetting, in which a random flow loses memory at prescribed equally spaced resetting times [13, 14]. The results achieved for both models agree well with the predictions of mixing-length theory. The first model is fully consistent with this theory; in the second type of model, the equations for the transport of magnetic field [13] and of a passive scalar [14] become integral over space, but the turbulent-diffusion rate for short correlation times remains the same as in mixing-length theory.

1063-7729/00/4410-0659$20.00

Since the prescribed resetting times violate time uniformity, the velocity field in resetting models is not a correlation function that depends on the time difference. S. A. Molchanov called our attention to the fact that this time uniformity can be restored if the memory-loss times are considered to be random Poisson rather than preassigned events. Such time restoration models can be solved exactly, but detailed computations have not been performed. The objective of the present study is to carry out such computations.

As a result of the transition from resetting models to time restoration models, the mean-field equation becomes even more complicated: it becomes an integral equation not only over space, but also over time. However, as we will see below, the transport preserves diffusion features. Surprisingly, the turbulent diffusion coefficient differs by a factor of two from the value predicted by mixing-length theory. We can illustrate this difference using a simple example. Consider two cases of particle random walk along a straight line. In the first case, the particle jumps at the times $t_1 = \tau$, $t_2 = 2\tau$, $t_3 = 3\tau\ldots$, a distance $V\tau$ to the left or right with equal probability one-half (this is the model used in mixing-length theory). In the second model, the times $t_1, t_2, \ldots$ correspond to a Poisson flow of events with mean time interval τ between events ($\lambda = 1/\tau$ is the parameter of the Poisson process). In both cases, the behavior of the particle over sufficiently long times can be described by a diffusion approximation, and direct calculation shows that, in the second case, the diffusion coefficient is twice that in the first case. The physical nature of this difference is that, in the second case, there exist comparatively long time intervals when the particle executes motion in a single direction. Although the probability of these time intervals is low, they contribute significantly to the diffusivity. This circumstance makes the considered effect resemble the phenomenon of intermittency [15], which is expected to be more prominent for the transport of higher moments. We stress that the distinction between the two models is unrelated to the distinction between t_n and $n\tau$, since the difference between these two quantities is of the order of $n^{1/2}$ and does not contribute to the turbulence diffusion coefficient. Moreover, the distinction exists even when the velocity is an arbitrarily distributed random quantity with zero mean (see Appendix).

Our objective in this work is to investigate this physical effect and, accordingly, we restrict our consideration to the simplest equation of transport of a passive scalar admixture in a statistically uniform and isotropic flow of incompressible fluid with time restoration. In other words, among all possible transport coefficients, we consider only those pertaining to turbulent diffusion. Similar effects certainly arise for other transport coefficients, such as those associated with the α effect. However, the corresponding calculations are more cumbersome, and we hope to realize them in a separate study.

2. VELOCITY FIELD AND BASIC EQUATIONS

Let us now specify the problem more exactly. We will consider in some probabilistic space a Poisson process $\eta(t)$ with parameter λ and a sequence of equally distributed steady-state velocity fields $\mathbf{v}^j(t, \mathbf{x})$, $j = 0, 1, 2\ldots, \eta$. The quantities η and $\mathbf{v}^j$ are assumed to be independent. $\tau(t)$ is the time of the last jump before time t. If there was no jump, $\tau(t) = 0$. We introduce the function $\sigma = \sigma(t) = t - \tau(t)$ and, finally, the velocity field $\mathbf{v}(t, \mathbf{x}) = \mathbf{v}^{\eta(t)}(\sigma(t), \mathbf{x})$. Direct calculation shows that the correlation function for this field takes the form

$$r_{nm}(t, s, \mathbf{x}, \mathbf{y}) = \hat{r}_{nm}(t - s, \mathbf{x}, \mathbf{y})e^{-\lambda|t-s|},$$

where $\hat{r}_{nm}$ is the correlation function for the fields v^j, i.e., the constructed field is uniform in time.

If we take $\hat{r} \sim \exp(-\hat{\lambda}t)$, then $r \sim \exp(\tilde{\lambda}t)$ and $\tilde{\lambda} = \hat{\lambda} + \lambda$. In other words, in the case of several memory-loss mechanisms, the reciprocals of the memory times add.

The evolution of the concentration of an admixture, $n(t, \mathbf{x})$, in the incompressible velocity field $\mathbf{v}(t, \mathbf{x})$ is described by the equation

$$\frac{\partial n}{\partial t} + \mathbf{v}\nabla n = D\Delta n, \tag{1}$$

where D is the coefficient of molecular diffusion.

The solution to the Cauchy problem for (1) is given by a Katz–Feynman formula [11]:

$$n(t, \mathbf{x}) = E_\xi\{n(s, \xi_t^x(s))\}. \tag{2}$$

Here, ξ_t^x is the quasi-Lagrangian trajectory of a particle that arrived at the point $\mathbf{x}$ at time t. This trajectory represents a solution to the integral equation

$$\xi_t^x(s) = \mathbf{x} - \int_s^t \mathbf{v}(u, \xi_t^x(u))du + \sqrt{2D}(\mathbf{w}(t) - \mathbf{w}(s)), \tag{3}$$
$$\xi_t^x(t) = \mathbf{x},$$

where $w(t)$ is a Wiener process. The symbol $E_\xi\{\cdot\}$ denotes averaging over all realizations of the random process ξ_t^x (here and below, E denotes averaging over the quantity indicated by the subscript). Thus, the expression $E_\xi\{n(s, \xi_t^s(s))\}$ is an integral equation over a space of functions with measure P_ξ induced by the process ξ_t^x. In this integral, it is helpful to make the substitution of variables (measures) [13] $\xi_t^x \longrightarrow \zeta_t^x$, where $\zeta_t^x(s) = x + \sqrt{2D}\,(\mathbf{w}(t) - \mathbf{w}(s))$. As a result, we obtain

$$n(t, \mathbf{x}) = E_\zeta\{n(s, \zeta_t^x(s))J(s, t, \zeta_t^x)\}, \tag{4}$$

where $J(s, t, \zeta_t^x)$ is the Jacobian of the transformation and E_ζ is the mean over the new measure. An explicit form of the function J was found in [13].

Owing to the Markovian property of the Wiener process, we have $E_\zeta\{\cdot\} = E_y\{E_\mu\{\cdot\}\}$, where E_μ is the average over the Wiener trajectory $\mu_t^{x,y}(s)$ connecting the points $(t, \mathbf{x})$ and $(s, \mathbf{y})$, and E_y is the average over the quantity $y = \zeta(s)$, which is normally distributed with mean $\mathbf{x}$ and variance $D(|t - s|)$. Finally, we obtain the formula

$$n(t, \mathbf{x}) = \int Q(t, s, \mathbf{x}, \mathbf{y}) \cdot n(s, \mathbf{y}) d^3\mathbf{y}, \tag{5}$$

with

$$Q(t, s, \mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{4\pi D|t-s|}} e^{\frac{-|x-y|^2}{4D|t-s|}} \cdot E_\mu\{J(s, t, \mu_t^{x,y})\}, \tag{6}$$

which is valid for any velocity field. However, this formula is useful for random velocity fields with abrupt disappearance of correlation at certain times, since it makes it possible to derive equations for the mean concentration without resorting to approximate hypotheses of correlation splitting.

3. EQUATION FOR THE MEAN CONCENTRATION

Let us formulate the problem in which we are interested. At the initial time ($t = 0$), an admixture with concentration $n_0(\mathbf{x})$ is introduced into a medium. Further, the admixture evolves in the random velocity field $\mathbf{v}(\mathbf{x})$. Thus, its concentration $n(t, \mathbf{x})$ is a random quantity. We are interested in the mean value of $n(t, \mathbf{x})$.

The equation for the mean concentration can be obtained by averaging equality (5) over all realizations of the velocity field. Note that the following factorization of the averaging scheme is possible: $E(\cdot) = E_\sigma(E_\tau(E_\nu(\cdot)))$, where E represents averaging over all realizations of the velocity field; E_ν averaging over all realizations of the velocity field having the given realization $\eta(t)$ of the Poisson process; E_τ averaging over all realizations of the Poisson process in which the last jump occurs at a time τ preceding the time t; and E_σ averaging over all times for the last jump before time t.

Let us average (5) over all realizations of the velocity field having the given Poisson trajectory. Substituting s for τ, we obtain

$$E_\nu(n(t, \mathbf{x})) = \int E_\nu(Q(t, \tau, \mathbf{x}, \mathbf{y}) \cdot n(\tau, \mathbf{y})) d^3\mathbf{y}.$$

Note that the quantity $Q(t, \tau, \mathbf{x}, \mathbf{y})$ is determined by the values $\mathbf{v}^{\eta(t)}$ and $n(\tau, \mathbf{y})$ is determined by the values $\mathbf{v}^j$ with $j = 0, 1, \ldots \eta(t) - 1$. Thus, the quantities $Q(t, \tau, \mathbf{x}, \mathbf{y})$ and $n(\tau, \mathbf{y})$ are independent of E_ν, and we have

$$E_\nu(n(t, \mathbf{x})) = \int E_\nu(Q(t, \tau, \mathbf{x}, \mathbf{y})) \cdot E_\nu(n(\tau, \mathbf{y})) d^3\mathbf{y}.$$

We now average this last equation over all Poisson trajectories making the last jump a given time before time t:

$$E_\tau(E_\nu(n(t, \mathbf{x}))) = \int E_\tau(E_\nu(Q(t, \tau, \mathbf{x}, \mathbf{y})) \cdot E_\nu(n(\tau, \mathbf{y}))) d^3\mathbf{y}.$$

Obviously, $Q(t, \tau, \mathbf{x}, \mathbf{y})$ is independent of what happened before time τ, and we find

$$E_\tau(E_\nu(n(t, \mathbf{x}))) = \int E_\nu(Q(t, \tau, \mathbf{x}, \mathbf{y})) \cdot E_\tau(E_\nu(n(\tau, \mathbf{y}))) d^3\mathbf{y}.$$

Let $P(\sigma, \mathbf{x}, \mathbf{y})$ represent averaging of $Q(\sigma, 0, \mathbf{x}, \mathbf{y})$ over any of the quantities $\mathbf{v}^j$. We then have

$$E_\tau(E_\nu(n(t, \mathbf{x}))) = \int P(\sigma, \mathbf{x}, \mathbf{y}) \cdot E_\tau(E_\nu(n(\tau, \mathbf{y}))) d^3\mathbf{y}.$$

The mean concentration N is determined by the relationship $N(t, \mathbf{x}) = E(n(t, \mathbf{x})|\mathcal{F}_{\geq t})$, where the symbol $\mathcal{F}_{\geq t}$ marks events occurring after time t, and the vertical bar denotes averaging, with the proviso that these events are fixed. In other words, $N(t, \mathbf{x})$ signifies averaging of the admixture concentration over all events occurring before time t. If t is a determinate moment in time, $N(t, \mathbf{x}) = E(n(t, \mathbf{x}))$; if, however, t is a random time, then $N(t, \mathbf{x})$ is a random quantity. Note that $E_\tau(E_\nu(n(\tau, \mathbf{y}))) = N(\tau, \mathbf{x})$. Therefore,

$$E_\tau(E_\nu(n(t, \mathbf{x}))) = \int P(\sigma, \mathbf{x}, \mathbf{y}) \cdot N(\tau, \mathbf{y}) d^3\mathbf{y}.$$

Since $\tau = t - \sigma$, averaging over all times for the last jump preceding time t gives

$$E_\sigma(E_\tau(E_\nu(n(t, \mathbf{x})))) = \int E_\sigma(P(\sigma, \mathbf{x}, \mathbf{y}) \cdot N(t - \sigma, \mathbf{y})) d^3\mathbf{y}.$$

Taking into account the averaging factorization and writing $E_\sigma(\cdot)$ in full, we can rewrite the last equality in the form

$$E(n(t, \mathbf{x})) = \iint P(\sigma, \mathbf{x}, \mathbf{y}) \cdot N(t - \sigma, \mathbf{y}) \cdot \varphi_t(\sigma) d\sigma d^3\mathbf{y}.$$

Here, t is a determinate instant of time and, therefore, $E(n(t, \mathbf{x})) = N(t, \mathbf{x})$, and φ_t is the probability density of $\tau(t)$. As a result, we obtain

$$N(t, \mathbf{x}) = \int_0^t \int P(\sigma, \mathbf{x}, \mathbf{y}) \cdot N(t - \sigma, \mathbf{y}) \cdot \lambda e^{-\lambda\sigma} d\sigma d^3\mathbf{y} + e^{-\lambda t} \int P(t, \mathbf{x}, \mathbf{y}) \cdot n_0(y) d^3\mathbf{y}. \tag{7}$$

The solution to this equation provides the solution of the problem stated at the beginning of this section.

Let us compare (7) with similar equations obtained in other exactly solved models for the velocity field. In a short-correlation approximation, the equation for the mean concentration is a differential equation [11], while we found for a velocity field with time resetting that it became an integral equation over space [14]. At the same time, (7) is an integral equation over both space and time.

Note that the resultant equation is slightly nonstationary due to the presence of the last term. The nonstationary character of the equation is associated with the transient process described by the term with $e^{-\lambda t}$. We can consider a slightly different formulation of the problem in which an admixture was introduced into a random flow in the distant past and we know the concentration distribution at some other time $t = 0$. In this case, the time uniformity in Eq. (7) is restored, and we have every reason to believe that this has only a small effect on the behavior of the admixture at large t. Below, we will use this obvious hypothesis extensively, although a rigorous proof has not been obtained, since this would require considerable analytical effort. Thus, we arrive at the equation:

$$N(t, \mathbf{x}) = \int_0^\infty\int P(\sigma, \mathbf{x}, \mathbf{y}) \cdot N(t-\sigma, \mathbf{y}) \cdot \lambda e^{-\lambda\sigma} d\sigma d^3\mathbf{y}, \quad (8)$$

$$N(0, \mathbf{x}) = N_0(\mathbf{x}),$$

which is valid for all real t, whereas Eq. (7) is valid only for nonnegative t. Below, we will not consider the properties of the transient process, and will investigate Eq. (8).

We assume the velocity field to be statistically uniform and isotropic, so that $P(\sigma, \mathbf{x}, \mathbf{y}) = P(\sigma, \mathbf{x} - \mathbf{y})$. Thus, after a Fourier transformation, (8) becomes an algebraic equation in the spatial variables:

$$N(t, \mathbf{k}) = \int_0^\infty P(\sigma, \mathbf{k}) \cdot N(t-\sigma, \mathbf{k}) \cdot \lambda e^{-\lambda\sigma} d\sigma. \quad (9)$$

Note that the function $P(\sigma, k)$ was already encountered in [13, 14], where it was shown that

$$P(t, \mathbf{k}) = E(e^{i\mathbf{k}\cdot\xi^{(t)}}). \quad (10)$$

Upon performing the Fourier transformation with respect to time, (9) also takes the form of an algebraic equation with respect to the time variable; i.e.,

$$N(\hat{t}, \mathbf{k}) = P(t, \mathbf{k})\widehat{\lambda e^{-\lambda t}}\Theta(t) \cdot N(\hat{t}, \mathbf{k}). \quad (11)$$

Here, $\Theta(t)$ is the theta function, and the symbol $\hat{}$ denotes a Fourier transform. In this equation, we cannot cancel out $N(\hat{t}, \mathbf{k})$ because $N(\hat{t}, \mathbf{k})$ is a generalized function.

By virtue of uniformity and isotropy,

$$E(\xi^{(t)}) = 0, \quad E(\xi_m^{(t)}\xi_n^{(t)}) = 2Wf(t)\delta_{mn}, \quad (12)$$

where f is the trajectory time correlation function.

The usefulness of Eq. (8) depends on the number of cases in which the kernel P takes a form suitable for further application. We turn now to the construction of specific versions of velocity fields with time restoration in which this kernel can be calculated explicitly.

4. SHORT RESTORATION TIME

Consider the case $\lambda \gg 1$ and $\lambda t \gg 1$. Then,

$$P(\sigma, \mathbf{k}) = 1 - k_m k_n D_{mn}\sigma + O(\sigma^{5/2}), \quad (13)$$

where $D_{mn} = \left(D + \frac{V^2}{2}\sigma\right)\delta_{mn}$ and $E(v_m(x)v_n(x)) = V^2\delta_{mn}$. Expanding (11) in terms of small λ and returning to the x representation, we obtain

$$\frac{\partial N}{\partial t} = \left(D + \frac{V^2}{\lambda}\right)\Delta N. \quad (14)$$

When comparing this result with estimates obtained using mixing-length theory, we should take into account that $V^2 = (v^*)^2/3$ and $\lambda^{-1} = t^*/2$. In other words, as in the example presented in the introduction, we derive a turbulent diffusion coefficient that is twice that obtained for mixing-length theory. Of course, comparing the parameters V and λ with the correlation rate and time in mixing-length theory is somewhat conditional. However, the result is twice the corresponding value obtained in a short-correlation approximation for a flow with resetting, and both results are initially expressed in terms of comparable variables.

5. VELOCITY FIELD WITH GAUSSIAN TRAJECTORIES

We will now consider velocity fields $\mathbf{v}^j$ whose trajectories ξ^t obey Gaussian statistics at any moment in time. This condition places strong restrictions on the set of fields under consideration, but enables calculation of the function P.

From (10), we obtain $P = \exp\left(-\mathbf{k}^2\frac{wf(t)}{2}\right)$, which gives an explicit, but somewhat awkward, form of the function $P(t, \mathbf{k})\widehat{\lambda e^{-\lambda t}}\Theta(t)$ in (11). This function will

become more compact if we assume in addition that $t_* \ll 1$ and $t_*\lambda \ll 1$. In this case,

$$P(t, \mathbf{k})\widehat{\lambda e^{-\lambda t}\Theta(t)} = \frac{\lambda\left(1 + \frac{wt_*}{2}\mathbf{k}^2\right)}{\lambda + w\mathbf{k}^2 + iu} + O(t_*^{3/2} + (t_*\lambda)^2). \quad (15)$$

For definiteness, we assume that

$$f(t) = \begin{cases} t^2/2t_*, & 0 \le t < t_* \\ t - t_*/2, & t_* \le t. \end{cases} \quad (16)$$

Our choice of the function f requires clarification. At sufficiently large t, $\xi^{(t)}$ is the sum of a large number of nearly independent quantities, and the relation between the variance $\xi_n^{(t)}$ and t is, therefore, linear, as in the case of Brownian motion. However, when t is small, we have $\xi_n^{(t)} \approx v_n t$ and $D(\xi_n^{(t)}) \approx t^2 D(v_n)$; i.e., the dependence becomes quadratic. For this reason, the function f is also quadratic before time t_*, and becomes linear after this time. Strictly speaking, at very small t, f again becomes linear, on account of molecular diffusion. However, we will neglect this effect to simplify the calculations.

Substituting (15) in (11), omitting high-order terms in small quantities, and returning to the coordinate representation, we obtain

$$\frac{\partial N}{\partial t} = w\left(1 - \frac{\lambda t_*}{2}\right)\Delta N. \quad (17)$$

This equation is identical to that inferred under similar assumptions in [14], although the diffusion coefficient is slightly smaller. The difference in the coefficients is due to the nonlinear character of f. Note that the coefficient of turbulent diffusion is now slightly smaller than the corresponding coefficient obtained in the model with time resetting [14]. This can be attributed to the fact that our choice of f artificially underestimates the probability of occurrence of large time intervals with high time correlations of the velocity field.

6. A WEAKLY ANISOTROPIC DISTRIBUTION OF THE ADMIXTURE

Assuming that the initial distribution of the admixture is sufficiently uniform, it is enough to calculate the function $P(\sigma, \mathbf{k})$ up to second-order terms in $\mathbf{k}$. We obtain from (10) and (12)

$$P(\sigma, \mathbf{k}) = 1 + ik_m E(\xi_m^{(\sigma)}) - \frac{1}{2}E(\xi_m^{(\sigma)}\xi_n^{(\sigma)}) = 1 - wf(\sigma)k_m k_n \delta_{mn},$$

$$N(t, \mathbf{x}) = \int_0^\infty N(t - \sigma, \mathbf{x})\lambda e^{-\lambda\sigma} d\sigma - w\Delta\int_0^\infty N(t - \sigma, \mathbf{x}) f(\sigma)\lambda e^{-\lambda\sigma} d\sigma. \quad (18)$$

We will search for a solution to (18) for the Fourier harmonics $N(t, \mathbf{x}) = e^{\gamma t + i\mathbf{kx}}$. In this case, Eq. (18) reduces to the dispersion relation

$$1 = \lambda\int_0^\infty e^{-(\gamma + \lambda)\sigma}(1 - k^2 wf(\sigma))d\sigma. \quad (19)$$

When solving this equation, we should retain terms not higher than k^2. Differentiating once with respect to k and setting $k = 0$, we obtain from (19) $\gamma(0) = \frac{\partial\gamma}{\partial k}(0) = 0$ and $\frac{\partial^2\gamma}{\partial k^2}(0) = -2w\lambda^2\int_0^\infty f(\sigma)e^{-\lambda\sigma}d\sigma$. Comparison with the dispersion relationship for the diffusion equation ($\gamma = -D_T k^2$) yields the final expression for the coefficient of turbulent diffusion

$$D_T = w\lambda^2\int_0^\infty f(\sigma)e^{-\lambda\sigma}d\sigma.$$

Obviously, $D_T > 0$.

7. DISCUSSION

Thus, we have revealed for the first time in the framework of exactly solved models the dependence of the coefficient of turbulent diffusion on the time-correlation properties of the system. Of course, if the vortex turnover time is comparable (to order of magnitude) to the memory time, we are only dealing with a change of a numerical factor. However, this factor can also be important for qualitative models of transport in specified objects.

Flow with time restoration bears little resemblance to laboratory turbulence that appears, for example, behind a grid. However, such flow resembles interstellar turbulence, which is probably maintained by supernova explosions. These explosions represent a good model for a Poisson flow of events, and can be treated as unrelated to the hydrodynamic memory loss mechanisms. The difference from models with regular time resetting is that the loss of memory in a medium with explosions occurs locally, near the explosion site, not in the entire system simultaneously. Of course, the memory time in such a medium, taken to be the mean time between supernova explosions in a given region, does

not necessarily coincide with the vortex turnover time (although the ratio of these time intervals—the so-called Strouhal number—does not differ greatly from unity [2]). Since supernovae are localized predominantly near the equatorial plane, the Strouhal number may depend on height above the Galactic equator. This dependence should be taken into account in models of transport processes in galaxies.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 00-02-17854, 00-02-16271). We are grateful to S.A. Molchanov, N. Kliorin, and I. Rogachevskiĭ for valuable discussions.

APPENDIX

Let V_i, $i = 0, 1, \ldots$ be a sequence of independent, equally distributed random variables with $EV_i = 0$ and τ be the resetting time, so that $t \in [n\tau; (n+1)\tau]$ and $v_{\text{ren}}(t) = V_n$. The displacement of a particle after n resetting events is $\xi_{\text{ren}} = \tau\sum_{i=0}^{n-1} V_i$, and its variance, which determines the diffusion coefficient, has the form $D\xi_{\text{ren}} = n\tau^2 DV_i$. Consider now a model with random resetting times. Let $\{\tau_i\}_{i=0}^{\infty}$ be a sequence of independent, exponentially distributed random variables with mean τ. The quantities V_i and τ_i are independent in total. The velocity field is determined as $v_{\text{rec}}(t) = V_i$ at $t \in [\tau_{i-1}; \tau_i]$, $\tau_{-1} = 0$. The particle displacement after n time restorations is $\xi_{\text{rec}} = \sum_{i=0}^{n-1} V_i\tau_i$ and

$$D\xi_{\text{rec}} = n(E(D(V_i\tau_i|\tau_i)) + D(E(V_i\tau_i|\tau_i)))$$

$$= n(DV_iE\tau_i^2 + (EV_i)^2 DV_i) = 2n\tau^2 DV_i.$$

Thus, we obtain $D\xi_{\text{rec}} = 2D\xi_{\text{ren}}$. These formulas can be compared to the results of mixing-length theory provided τ is identified with $2t^*$, and DV_i, with $(v^*)^2/3$.

REFERENCES

1. F. Krauze and K.-Kh. Radler, *Field Magnetohydrodynamics and Dynamo Theory* (Akademie, Berlin, 1980; Mir, Moscow, 1984).
2. A. A. Ruzmaĭkin, D. D. Sokolov, and A. M. Shukurov, *Galactic Magnetic Fields* (Nauka, Moscow, 1988).
3. S. I. Vainshtein and F. Cattaneo, Astrophys. J. **393**, 165 (1992).
4. G. Rüdiger and L. L. Kitchatinov, Astron. Astrophys. **269**, 581 (1993).
5. R. Beck, A. Brandenburg, D. Moss, *et al.*, Ann. Rev. Astron. Astrophys. **34**, 155 (1996).
6. M. J. Korpi, A. Brandenburg, A. Shukurov, *et al.*, Astrophys. J. Lett. **514**, L99 (1999).
7. A. Z. Dolginov and N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **93**, 159 (1987) [Sov. Phys. JETP **66**, 90 (1987)].
8. L. L. Kitchatinov, V. V. Pipin, and G. Rüdiger, Astron. Nachr. **315**, 157 (1994).
9. V. I. Arnol'd, Ya. B. Zel'dovich, A. A. Ruzmaĭkin, and D. D. Sokolov, Zh. Éksp. Teor. Fiz. **81**, 2052 (1981) [Sov. Phys. JETP **54**, 1083 (1981)].
10. Ya. B. Zeldovich, A. A. Ruzmaĭkin, S. A. Molchanov, and D. D. Sokoloff, J. Fluid Mech. **144**, 1 (1984).
11. Ya. B. Zeldovich, A. A. Ruzmaikin, S. A. Molchanov, and D. D. Sokoloff, Sov. Sci. Rev., Sect. C: Math. Phys. **7**, 1 (1988).
12. S. A. Molchanov, A. A. Ruzmaĭkin, and D. D. Sokoloff, Magn. Gidrodin., No. 4, 67 (1983).
13. P. Dittrich, S. A. Molchanov, D. D. Sokoloff, and A. A. Ruzmaĭkin, Astron. Nachr. **305** (3), 119 (1984).
14. T. Elperin, N. Kliorin, I. Rogachevskiĭ, and D. Sokoloff, Phys. Rev. E. **61** (2) (2000).
15. Ya. B. Zel'dovich, S. A. Molchanov, A. A. Ruzmaĭkin, and D. D. Sokolov, Usp. Fiz. Nauk. **152**, 3 (1987) [Sov. Phys. Usp. **30**, 353 (1987)].

Translated by A. Kozlenkov

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 665–684. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 750–771.

Globular Cluster Subsystems in the Galaxy

T. V. Borkova and V. A. Marsakov

Rostov State University, pr. Stacki 194, Rostov-on-Don, 344104 Russia

Received December 4, 1998

Abstract—Data from the literature are used to construct a homogeneous catalog of fundamental astrophysical parameters for 145 globular clusters of the Milky Way Galaxy. The catalog is used to analyze the relationships between chemical composition, horizontal-branch morphology, spatial location, orbital elements, age, and other physical parameters of the clusters. The overall globular-cluster population is divided by a gap in the metallicity function at [Fe/H] = –1.0 into two discrete groups with well-defined maxima at [Fe/H] = $–1.60 \pm 0.03$ and $–0.60 \pm 0.04$. The mean spatial–kinematic parameters and their dispersions change abruptly when the metallicity crosses this boundary. Metal-poor clusters occupy a more or less spherical region and are concentrated toward the Galactic center. Metal-rich clusters (the thick disk subsystem), which are far fewer in number, are concentrated toward both the Galactic center and the Galactic plane. This subsystem rotates with an average velocity of $V_{\rm rot} = 165 \pm 28$ km/s and has a very steep negative vertical metallicity gradient and a negligible radial gradient. It is, on average, the youngest group, and consists exclusively of clusters with extremely red horizontal branches. The population of spherical-subsystem clusters is also inhomogeneous and, in turn, breaks up into at least two groups according to horizontal-branch morphology. Clusters with extremely blue horizontal branches occupy a spherical volume of radius ~9 kpc, have high rotational velocities ($V_{\rm rot} = 77 \pm 33$ km/s), have substantial and equal negative radial and vertical metallicity gradients, and are, on average, the oldest group (the old-halo subsystem). The vast majority of clusters with intermediate-type horizontal branches occupy a more or less spherical volume ≈18 kpc in radius, which is slightly flattened perpendicular to the Z direction and makes an angle of ≈30° to the X-axis. On average, this population is somewhat younger than the old-halo clusters (the young-halo subsystem), and exhibits approximately the same metallicity gradients as the old halo. As a result, since their Galactocentric distance and distance from the Galactic plane are the same, the young-halo clusters have metallicities that are, on average, Δ[Fe/H] ≈ 0.3 higher than those for old-halo clusters. The young-halo subsystem, which apparently consists of objects captured by the Galaxy at various times, contains many clusters with retrograde orbits, so that its rotational velocity is low and has large errors, $V_{\rm rot} = -23 \pm 54$ km/s. Typical parameters are derived for all the subsystems, and the mean characteristics of their member globular clusters are determined. The thick disk has a different nature than both the old and young halos. A scenario for Galactic evolution is proposed based on the assumption that only the thick-disk and old-halo subsystems are genetically associated with the Galaxy. The age distributions of these two subsystems do not overlap. It is argued that heavy-element enrichment and the collapse of the proto-Galactic medium occurred mainly in the period between the formation of the old-halo and thick-disk subsystems. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Quite recently, globular clusters (GCs) were considered to form a homogeneous group and to be typical representatives of the spherical component of the Milky Way. However, a large amount of recently published observational material has demonstrated a substantial scatter in the physical and chemical parameters of clusters, and shown that the distributions of these parameters are discrete. This suggests the existence of several populations of GCs belonging to different subsystems of the Galaxy. Even the earliest metallicity functions revealed a gap near [Fe/H] ≈ –1.0, which divides the GC population into two discrete groups: a metal-poor, spherically symmetric, slowly rotating halo subsystem and a metal-rich, rather rapidly rotating, thick-disk subsystem [1, 2]. Halo GCs were further shown to separate into two groups with different horizontal-branch (HB) morphologies. These subgroups, whose distributions are both spherical, differ in their kinematics and the spatial volume they occupy [3]. Halo clusters, which have redder HBs for a given metallicity, are mostly located outside the solar circle and have a large velocity dispersion, lower rotational velocities, and smaller ages than clusters with blue HBs, which are concentrated inside the solar circle [4]. This difference can be explained if the old-halo subsystem formed simultaneously with the entire Galaxy, whereas the young-halo clusters were captured from intergalactic space during later evolutionary stages [3].

The aim of this paper is to analyze relationships between the physical, chemical, and spatial–kinematic parameters of GCs both for the Galaxy as a whole and within each subsystem, and to determine the characteristic parameters of the subsystems. This requires, first and foremost, a homogeneous catalog of fundamental GC parameters.

1063-7729/00/4410-0665$20.00 © 2000 MAIK "Nauka/Interperiodica"

2. THE CATALOG OF GLOBULAR CLUSTERS

Our catalog is based on the computer-readable version of the compiled catalog of Harris [5], which gives measured quantities for 147 Galactic GCs. These data are complete through May 15, 1997. We adopted most parameters directly from [5] and computed some using data from this same catalog. We further added some fundamental parameters that are missing from [5], taking them from other sources.

We adopted the positions of the clusters in the Galaxy from Harris [5], who used the horizontal-branch magnitude averaged over several sources as his main distance indicator. We transformed the Galactocentric coordinates given in [5] into Galactic coordinates X, Y, and Z (for a Galactocentric distance of the Sun $R_G = 8$ kpc), and computed $\cos\psi$, where ψ is the angle between the GC line of sight and the vector of rotation about the Z-axis, using the formula

$$\cos\psi = \frac{R_G Y}{R_\Theta \sqrt{X^2 + Y^2}}.$$

Here, R_Θ is the heliocentric distance of the GC.

The radial velocities (V_r) in [5] were derived by averaging the data for a large number of sources using weights inversely proportional to the errors, which were ~1 km/s for most of the sources. We also computed V_S, the GC line-of-sight velocities relative to an observer at rest at the position of the Sun, using the formula

$$V_S = V_0 + V\cos A,$$

where V_0 is the radial velocity corrected for the solar motion relative to the local centroid from [5], V = 225 km/s is the velocity of circular motion of the local centroid of the Sun, and A is the angle between the apex of the circular motion of the local centroid of the Sun and the direction to the GC ($\cos A = Y/R$). See [6] for a detailed description of the angle and velocity computations.

One of the most important parameters of a cluster is its metallicity. In his catalog, Harris [5] compiled all published GC metallicities reduced to the [Fe/H] scale of [7] and averaged them with equal weights. The mean [Fe/H] values in the compiled catalog have relatively high internal accuracies, due to the large number of sources considered (>40) and the exclusive use of spectroscopic metallicity determinations and well calibrated color–magnitude diagrams. The metallicity scale [7] is somewhat nonlinear, resulting in overestimated metallicities for the most metal-rich clusters [10]; however, only the relative [Fe/H] values are important to us here.

The horizontal branch can be used not only to determine the cluster distance, but also to obtain information about the conditions under which GCs form and evolve. In particular, Mironov and Samus' [11] subdivided all clusters into two groups based on their HB morphologies and found them to differ distinctly in their metallicities and spatial–kinematic characteristics. On the other hand, according to Oosterhoff, GCs divide into two distinct groups separated by a gap in the period distribution for their typical horizontal-branch representatives—*RR* Lyr variables. The periods of *RR* Lyrae in clusters are closely related to the HB morphology, allowing clusters to be conveniently characterized by the parameter $(B - R)/(B + V + R)$, where B, V, and R are the numbers of stars in the blue end of the HB, the Hertzsprung gap, and the red end, respectively. Harris [5] computed this parameter by simply averaging the data for several sources.

Age t is one of the most uncertain parameters, and Harris [5] does not give it in his compiled catalog. The recently published Hipparcos catalog contains high-precision stellar parallaxes based on satellite measurements, and even a first-level analysis of these data requires substantial revision of GC distances. As a result, the ages of even the most metal-poor (i.e., oldest) clusters do not exceed ~10 Gyr (see, e.g., [12]). However, we adopted the old scale here, because the refinement of the age-scale zero point based on the new data is probably now only in its initial stage, and we are primarily interested in relative parameters.

Accurate relative GC ages have recently been published in a number of studies. To be able to use age data for as many clusters as possible, we reduced all these age lists to a unified scale and computed weighted average age estimates by assigning weights both to each source and to each individual age determination. We used the two-tiered iteration procedure suggested by Hauck and Mermilliod [13], assigning lower weights to age determinations that differ strongly from the initially computed mean. We used homogeneous relative ages for 36 GCs from [14] as our basic scale, where the mean age of metal-poor clusters was assumed to be 15 Gyr. We then used a least squares method to reduce the age scales of the nine most extensive lists (containing 26 [15], 12 [16], 31 [17], 32 [18], 14 [19], 40 [20], 43 [21], 36 [22], and 25 [23] GCs) to the reference age scale. We included age determinations for single clusters from other studies only if their theoretical isochrones coincided with those adopted in one of the papers listed above. We used a total of 47 sources (not given in the references) and 336 individual age determinations, and derived weighted average estimates for a total of 63 GCs. The resulting ages had an internal accuracy of $\sigma_t \approx 0.89 \pm 0.03$ Gyr.

The central concentrations $C = \log(r_t/r_c)$, where r_t and r_c are the tidal radius and the measured core radius, respectively, were taken from [5]. We transformed the cluster radii in angular units, which Harris [5] derived by averaging published angular measurements, into linear radii in pc (r_h) using the heliocentric cluster distances given in the same catalog. We adopted the central cluster densities $\log\rho_0$ from [24] and, for clusters absent from this source, from [25]. Both papers were written by the same team of authors, but they report dif-

ferent densities and cluster lists. We used the more recent paper as our basic source.

We took the orbital elements for 25 GCs from [26]. The cluster radial velocities and distances used in [26] differ somewhat from those adopted here, but only slightly. The high accuracy of the distances and proper motions in [26] is demonstrated, in particular, by the fact that the mean rotational velocities of the GC subsystems derived from the radial velocities alone agree well with those derived from the full velocity vectors, as pointed out by Douphole *et al.* [26]. We subjected these orbital elements to a statistical test. More than half of the GCs analyzed in [26] are located near computed apogalactic orbital radii, in full consistency with the theoretical phase distribution of GC orbital locations. We also found the minimum orbital radii from [26] to agree satisfactorily with the perigalactic distances estimated in [27] from the tidal criterion ($r = 0.55 \pm 0.17$). These orbital elements are thus suitable for analysis of the properties of GC populations.

We derived the cluster masses from the integrated absolute magnitudes from [5] assuming $M/L_V = 3$ [28], where the mass M and luminosity L_V are in solar units. Table 1 lists some of the above parameters for 145 GCs with known distances.

3. GLOBULAR CLUSTERS OF THE DISK AND HALO

Metallicity Function

Figure 1 shows the distribution of heavy-element abundance (i.e., the metallicity function) for all the GCs from Table 1. The solid curve shows an approximation of the histogram using a superposition of two Gaussian curves with parameters estimated using a maximum-likelihood method. The probability that we would be wrong to discard the null hypothesis that the distribution can be described by a single Gaussian in favor of a fit using the superposition of two Gaussians is $\ll 1\%$. Thus, the entire GC population can be divided into two metallicity groups, with maxima at [Fe/H] = -0.60 ± 0.04 and -1.60 ± 0.03 and equal dispersions $\sigma \approx 0.30 \pm 0.03$ separated by a well-defined gap at [Fe/H] = –1.0. When fitting the distribution, we ignored an outlier cluster with [Fe/H] = +0.22. Our gap position is shifted by Δ[Fe/H] ≈ -0.2 toward lower metallicities compared to Zinn's [2] result, which has usually been used to determine whether a particular cluster belongs to the halo or thick-disk population (see, e.g., [29]). (Note that the same metallicity scale is used in both cases.) The breakdown of the GC population into two subsystems separated by [Fe/H] = –1.0 is further supported by the fact that diagrams depicting the spatial locations and kinematics of the clusters show each characteristic parameter to have a discontinuity precisely at this (or somewhat lower) metallicity (see Fig. 5 below).

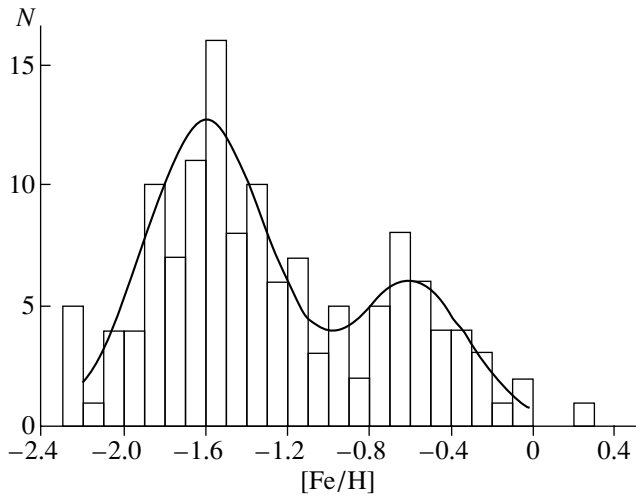


Fig. 1. Globular-cluster metallicity function. The curve shows an approximation by a superposition of two Gaussians.

Spatial Distribution

Figure 2 shows the distributions of metal-rich and metal-poor GC groups projected onto the *XY* and *YZ* planes. (The figures do not show six clusters with Galactocentric distances exceeding 60 kpc.) The two subsystems can easily be seen to differ strongly in both the volume and shape of the domain they occupy. The metal-rich group, which is much smaller in size, is concentrated toward both the center and the Galactic plane, and its shape can be very roughly described as an ellipsoid of revolution flattened along the *Z* coordinate. The envelope of its *XY* projection forms a circle ~7 kpc in radius, and the envelope of the *YZ* projection forms an ellipse with a *Z* semimajor axis of ≈3 kpc. A comparison of these parameters with those of high-velocity field stars suggests that this GC group belongs to the thick disk. (The differences in the sizes in the *X* and *Y* directions are probably due to the poor statistics for this group.)

The squares in Figs. 2a and 2b denote clusters that are far from the upper envelopes. The name of the cluster is given near each square. A more detailed discussion of these clusters is given below. We did not use these outlier clusters when determining the parameters of the thick disk. Similarly, the circles in Figs. 2c and 2d indicate halo clusters lying outside the circular envelope. However, we did not exclude these clusters when determining the characteristic parameters of the halo and Galaxy as a whole. To characterize the subsystem sizes, we used scale lengths and scale heights in (X_0, Y_0, Z_0), equal to the Galactocentric distances along the corresponding coordinates over which the cluster density decreases by a factor of *e*. The corresponding quantities are listed in Table 2.

Table 1. Master catalog of fundamental GC parameters

NGC	[Fe/H]	$\frac{B-R}{B+V+R}$	R, kpc	V_S, km/s	$\cos\psi$	t, Gyr	$\log(M/M_\odot)$	R_a, kpc	Z_{max}, kpc	e	r_h^*	$\log\rho_0$	C	Subsystem
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
104	−0.76	−0.99	7.3	−154.3	−0.687	14.64	6.16	7	3	0.07	4.76	5.14	2.04	1
288	−1.24	0.98	11.4	52.8	0.000	15.08	5.03	12	9	0.51	5.20	1.80	0.96	2
362	−1.16	−0.87	9.2	83.6	−0.674	13.78	5.75	10	8	0.80	1.92	4.75	1.94	3
1261	−1.35	−0.71	17.9	−81.6	−0.389	15.5	5.51				2.47	2.98	1.27	3
Pal 1	−0.80	−1.00	15.9		0.372		3.16				1.23		1.50	3(1)
AM1	−1.80	−0.93	120.5	−43.5	−0.064	10.74	4.27						1.23	3
ERID	−1.46	−1.00	85.8	−143.2	−0.055	11.88	4.37				2.96	0.33	1.10	3
Pal 2	−1.30		34.8	−106.2	0.038		5.59				2.86		1.45	3
1851	−1.26	−0.36	16.8	139.5	−0.388	13.17	5.75				1.49	5.16	2.24	3
1904	−1.54	−0.89	18.5	47.5	−0.294	14.05	5.53				2.34	4.01	1.72	3(2)
2298	−1.85	0.93	15.6	−62.8	−0.457	15.23	4.91				2.09	3.48	1.28	3(2)
2419	−2.12	0.86	89.6	−27.7	−0.001	14.90	6.22				5.45	1.50	1.40	3
Pyxis	−1.20	−1.00	40.5		−0.195	12.82	4.68				6.89		0.65	3
2808	−1.37	−0.49	10.9	−130.8	−0.711	14.44	6.15				1.91	4.63	1.77	3
E3	−0.80		7.6		−0.938		3.5				3.47		0.75	1
Pal 3	−1.66	−0.82	92.6	−67.1	−0.073	13.42	4.65				5.13		1.00	3
3201	−1.48	0.08	8.9	266.1	−0.881	13.87	5.38				4.98	2.63	1.31	3
Pal 4	−1.48	−1.00	102.2	50.6	−0.025	13.92	4.74				4.43	−0.24	0.78	3
4147	−1.83	0.55	20.9	139.9	−0.169	15.49	4.85	52	46	0.59	1.53	3.58	1.80	3
4372	−2.09	1.00	6.9	−122.0	−0.980		5.42				7.10	2.18	1.30	2
Rup 106	−1.80	−0.82	18.0	−234.6	−0.385	12.33	4.94				4.11	1.22	0.70	3
4590	−2.06	0.17	10.0	−251.7	−0.692	14.52	5.33				4.05	2.52	1.64	3
4833	−1.79	0.93	6.9	11.3	−0.969		5.59				4.82	3.05	1.25	2
5024	−2.07	0.76	18.9	−89.3	−0.122	15.22	5.91	26	25	0.54	3.92	3.53	1.78	3
5053	−2.29	0.52	16.8	34.4	−0.119	15.32	5.08				11.59	0.51	0.82	3
5139	−1.62	0.89[1]	6.3	64.5	−0.966	16.43	6.51	7	3	0.63	7.77		1.24	2
5272	−1.57	0.08	11.9	−109.1	0.158	14.52	5.95	13	12	0.48	2.91	3.54	1.85	3
5286	−1.67	0.80	8.2	−106.1	−0.742		5.83				1.86	4.19	1.46	3
Am 4	−2.00		24.8		−0.226		3.03				1.87		0.50	3
5466	−2.22	0.58	16.9	161.4	0.275	15.94	5.23	52	50	0.75	7.54	0.68	1.43	3
5634	−1.82		21.3	−80.0	−0.173		5.49				2.24		1.60	3(R)
5694	−1.86	1.00	28.3	−232.3	−0.148		5.51				1.58	4.01	1.84	3(2)
I4499	−1.60	0.11	15.4		−0.428	13.11	5.32				5.29	1.50	1.11	2
5824	−1.85	0.79	25.1	−127.4	−0.154		5.93				1.66	4.65	2.45	3
Pal 5	−1.38	−0.40	18.0	−40.1	0.011	13.45	4.46	18	18	0.36	11.58	−0.77	0.74	3
5897	−1.80	0.86	7.6	54.3	−0.498	15.44	5.31				6.19	1.32	0.79	2
5904	−1.29	0.31	6.1	74.3	0.122	14.38	5.91	13	33	0.83	4.69	3.92	1.87	3
5927	−0.37	−1.00	4.5	−235.2	−0.984	9.14	5.53				2.57		1.60	1
5946	−1.38		7.1	3.3	−0.612		5.43				1.99		2.50	2(C)
BH 176		−1.00	9.4		−0.448		4.13							
5986	−1.67	0.97	4.7	10.7	−0.741		5.76				2.77	3.24	1.22	2
Lynga 7	−0.62	−1.00	4.2	−105.0	−0.998		–							1
Pal 14	−1.52	−1.00	67.3	171.8	0.061	12.27	4.28				8.04	0.05	0.72	3
6093	−1.62	0.93	3.1	−6.8	−0.912		5.58				1.58	4.79	1.95	2
6121	−1.20	−0.06	6.0	46.6	−0.204	15.05	5.27	6	3	0.84	4.45	3.91	1.59	3
6101	−1.82	0.84	11.0	214.9	−0.510	14.21	5.15				5.47	1.57	0.80	3
6144	−1.73	1.00	3.4	169.1	−0.509		5.21				4.24		1.55	2
6139	−1.65		3.9	−53.1	−0.650		5.77				2.19		1.80	2–3
Ter 3			18.9		−0.110		4.85				5.50		0.70	3(R)
6171	−1.04	−0.73	3.3	−8.7	0.193	14.80	5.24	3	3	0.22	5.58	3.44	1.51	3

Table 1. (Contd.)

NGC	[Fe/H]	$\frac{B-R}{B+V+R}$	R, kpc	V_S, km/s	$\cos\psi$	t, Gyr	$\log(M/M_\odot)$	R_a, kpc	Z_{max}, kpc	e	r_h^*	$\log\rho_0$	C	Sub-system
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ESO 452			2.0		−0.897		3.99							
6205	−1.54	0.97	8.3	−85.7	0.745	15.10	5.81	21	21	0.59	3.24	3.96	1.49	2
6229	−1.44	0.24	28.7	24.6	0.272	14.06	5.60				1.65	3.46	1.61	3
6218	−1.48	0.97	4.6	27.1	0.473	13.69	5.32	6	5	0.38	3.85	3.26	1.38	2
6235	−1.40	0.89	2.7	94.8	−0.102		4.85				2.15		1.33	2
6254	−1.52	0.98	4.6	144.4	0.446	15.30	5.38	5	5	0.28	3.09	3.50	1.40	2
6256	−0.70		2.3		−0.755		4.87				2.13		2.50	1
Pal 15	−1.90	1.00	36.9	150.0	0.073		4.58				6.57		0.60	3(2)
6266	−1.29	0.32	1.8	80.7	−0.561		6.07				2.62	5.34	1.70	3
6273	−1.68		1.5	133.7	−0.729		6.02				3.00		1.53	3
6284	−1.32		6.6	34.6	−0.037		5.54				2.43	4.65	2.50	2(C)
6287	−2.05	0.98	1.6	−195.9	0.072		5.25				1.79	3.99	1.60	2
6293	−1.92	0.90	1.5	−97.3	−0.413		5.50				2.22	6.01	2.50	2
6304	−0.59	−1.00	2.2	−113.3	−0.277		5.32				2.84		1.80	1
6316	−0.55	−1.00	3.7	71.0	−0.113		5.84				1.98		1.55	1
6341	−2.29	0.91	9.5	64.8	0.735	16.16	5.67	11	5	0.86	2.55	4.38	1.81	2
6325	−1.17		1.9	18.4	0.102		5.33				2.37		2.50	2(C)
6333	−1.72	0.87	1.7	262.9	0.953		5.31				2.25	3.58	1.15	2
6342	−0.65	−1.00	1.9	112.1	0.557		5.01				2.18		2.50	1
6356	−0.50	−1.00	7.0	65.5	0.142		5.80				2.33		1.54	1
6355	−1.50		1.2	−167.6	−0.063		5.39				1.91		2.50	2(C)
6352	−0.70	−1.00	3.3	−186.3	−0.777	13.62	4.98				3.89		1.10	1
IC 1257	−1.70	1.00	17.3		0.136		4.85							3(2)
Ter 2	−0.25	−1.00	1.6	104.6	−0.322		4.41				3.86		2.50	1
6366	−0.82	−0.97	4.9	−39.5	0.501	12.79	4.70				4.11	1.64	0.92	1
Ter 4	−1.60		1.1	−55.8	−0.518		4.83							2–3
HP1	−1.50		0.9	53.2	−0.412		4.97				6.84		2.50	2(C)
6362	−1.06	−0.58	5.1	−132.8	−0.946	14.52	5.18				4.91	2.23	1.10	3
LILL	0.22	−1.00	2.6	41.2	−0.273		5.44				1.20		2.30	1
Ter 1	−0.35		1.5	35.6	−0.221		3.71				8.01		2.50	1
Ton 2	−0.50		1.4		−0.998		4.85				2.50		1.30	1
6388	−0.60		4.4	32.2	−0.474		6.32				1.87		1.70	1
6402	−1.39	0.65	3.8	27.6	0.918		6.00				3.12	3.46	1.60	3
6401	−1.12		0.9	−40.0	0.689		5.44				4.30		1.69	2–3
6397	−1.95	0.98	6.0	−58.6	−0.481	16.52	5.04	6	2	0.35	2.84	5.70	2.50	2
Pal 6	−0.10	−1.00	1.4	220.2	0.219		5.34				2.26		1.10	1
6426	−2.26	0.53	13.8	−45.0	0.286		5.05				3.52		1.70	3
Djorg 1			1.4		−0.36		5.01				3.14		1.50	
Ter 5	−0.28	−1.00	0.6	−67.8	0.999		5.55				1.93		1.74	1
6440	−0.34	−1.00	1.2	−36.4	0.996		5.89				1.35		1.70	1
6441	−0.53		2.1	1.6	−0.464		6.08				1.64		1.85	1
Ter 6	−0.65	−1.00	0.6	130.2	−0.396		5.15				0.99		2.50	1
6453	−1.53		3.1	−91.3	−0.201		5.21				1.01		2.50	2(C)
UKS1	−1.20		0.8		0.836		4.88				1.94		2.10	2–3
6496	−0.64	−1.00	4.4	−152.0	−0.413	10.27	5.28				5.24		0.70	1
Ter 9	−1.00		0.6	83.8	0.935		3.96				1.79		2.50	1
Djorg 2			5.8		0.066		5.97				2.54		1.50	
6517	−1.37		4.1	47.9	0.671		5.70				1.65		1.82	2–3
Ter 10	−0.70		0.8		0.826		5.52							1
6522	−1.52	0.71	1.1	−1.0	0.139		5.41				2.26	5.53	2.50	3

Table 1. (Contd.)

NGC	[Fe/H]	$\frac{B-R}{B+V+R}$	R, kpc	V_S, km/s	$\cos\psi$	t, Gyr	$\log(M/M_\odot)$	R_a, kpc	Z_{max}, kpc	e	r_h^*	$\log\rho_0$	C	Subsystem
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6535	−1.80	1.00	3.9	−99.2	0.977	15.71	5.28				1.65	2.25	1.30	3
6528	−0.17	−1.00	0.8	199.4	0.249	14.00	5.08				0.96		2.29	1
6539	−0.66	−1.00	3.0	47.7	0.985		5.71				3.86		1.60	1
6540	−1.00		4.5		0.101		4.53				0.37		2.50	1
6544	−1.56	1.00	5.5	6.7	0.147		5.01				2.30		1.63	2
6541	−1.83	1.00	2.1	−190.4	−0.910		5.78				2.66		2.00	2
6553	−0.25	−1.00	3.3	24.9	0.218	14.00	5.50				2.77		1.17	1
6558	−1.44		1.8	−133.3	0.017		4.84				3.35		2.50	2(C)
IC 1276	−0.75[2]		3.6		0.848		5.35				5.90		1.29	1
6569	−0.86		1.1	−16.7	0.149		5.54				3.19		1.27	1
6584	−1.49	−0.15	6.5	161.3	−0.433	14.50	5.46				2.37	3.19	1.20	3
6624	−0.42	−1.00	1.2	74.5	0.911	11.25	5.39				1.90		2.50	1
6626	−1.45	0.90	2.5	58.2	0.432	19.10	5.72	2	1	0.25	3.06	4.71	1.67	2
6638	−0.99	−0.30	1.5	59.4	0.974		5.12				1.56	4.02	1.40	1
6637	−0.71	−1.00	1.5	55.8	0.942	11.11	5.40				1.96		1.39	1
6642	−1.35		1.6	−8.1	0.966		5.02				1.66		1.99	2–3
6652	−0.96	−1.00	2.2	−96.8	0.169	12.37	5.02				1.64		1.80	1
6656	−1.64	0.91	5.0	−99.7	0.278	15.18	5.79	10	2	0.55	4.80	3.67	1.31	2
Pal 8	−0.48	−1.00	5.2	22.9	0.391		4.60				1.65		1.53	1
6681	−1.51	0.96	2.0	238.6	0.605		5.23				2.26	5.56	2.50	2
6712	−1.01	−0.64	3.5	1.3	0.985		5.39	7	4	0.92	2.92	3.09	0.90	3
6715	−1.59	0.87	18.6	172.5	0.043		6.39				2.06		1.84	3(2)
6717	−1.29	0.98	2.4	82.8	0.891		4.66				1.49	4.60	2.07	2
6723	−1.12	−0.08	2.6	−86.4	0.044		5.53				3.89	2.71	1.05	3
6749	−1.60	1.00	4.9		0.967		5.07				2.51		0.83	2
6752	−1.55	1.00	5.3	−103.5	−0.578	15.18	5.48				3.80	4.88	2.50	2
6760	−0.52	−1.00	4.8	118.5	0.983		5.55				4.85		1.59	1
6779	−1.94	0.98	9.5	76.5	0.750		5.34	33	5	0.81	4.85		1.37	2
Ter 7	−0.58	−1.00	15.8	186.3	0.033	9.32	4.41				3.83		1.08	3(1)
Pal 10	−0.10	−1.00	6.4		0.989		4.71				1.96		0.58	1
Arp 2	−1.76	0.86	20.4	154.7	0.062	12.93	4.51				8.26	−0.35	0.90	3(2)
6809	−1.81	0.87	3.9	214.3	0.343	15.43	5.41				5.47	2.12	0.76	2
Ter 8	−1.87		18.5	157.8	0.048	15.87	4.43				4.15		0.60	3(C, R)
Pal 11	−0.39		7.6	57.8	0.597		5.11				4.35		0.69	3(1)
6838	−0.73	−1.00	6.7	177.9	0.993	14.56	4.61	7	0	0.19	2.65	3.29	1.15	1
6864	−1.32	−0.42	12.4	−110.4	0.264		5.73				1.66	4.53	1.88	3
6934	−1.54	0.25	12.4	−232.5	0.525		5.39	32	22	0.49	1.93	3.43	1.53	3
6981	−1.54	0.14	12.7	−228.1	0.436	13.07	5.21				2.97	2.26	1.23	3
7006	−1.68	−0.28	38.0	−177.1	0.190	14.27	5.47				2.00	2.42	1.42	3
7078	−2.22	0.67	10.3	82.2	0.699	15.43	6.05	37	30	0.60	2.79	5.30	2.50	3
7089	−1.62	0.96	10.3	148.3	0.666	15.25	6.00	28	24	0.67	2.58	3.89	1.80	2
7099	−2.12	0.89	7.0	−109.8	0.620	15.46	5.36				2.66	5.07	2.50	2
Pal 12	−0.93	−1.00	15.5	108.7	0.391	11.34	4.18				4.55	0.68	1.94	3(1)
Pal 13	−1.65	−0.20	27.2	139.8	0.286	13.27	3.79				1.94		0.66	3
7492	−1.51	0.81	24.3	−127.2	0.314	15.79	4.70				5.04	1.25	1.00	3

[1]Adopted from [15].
[2]Adopted from [9].

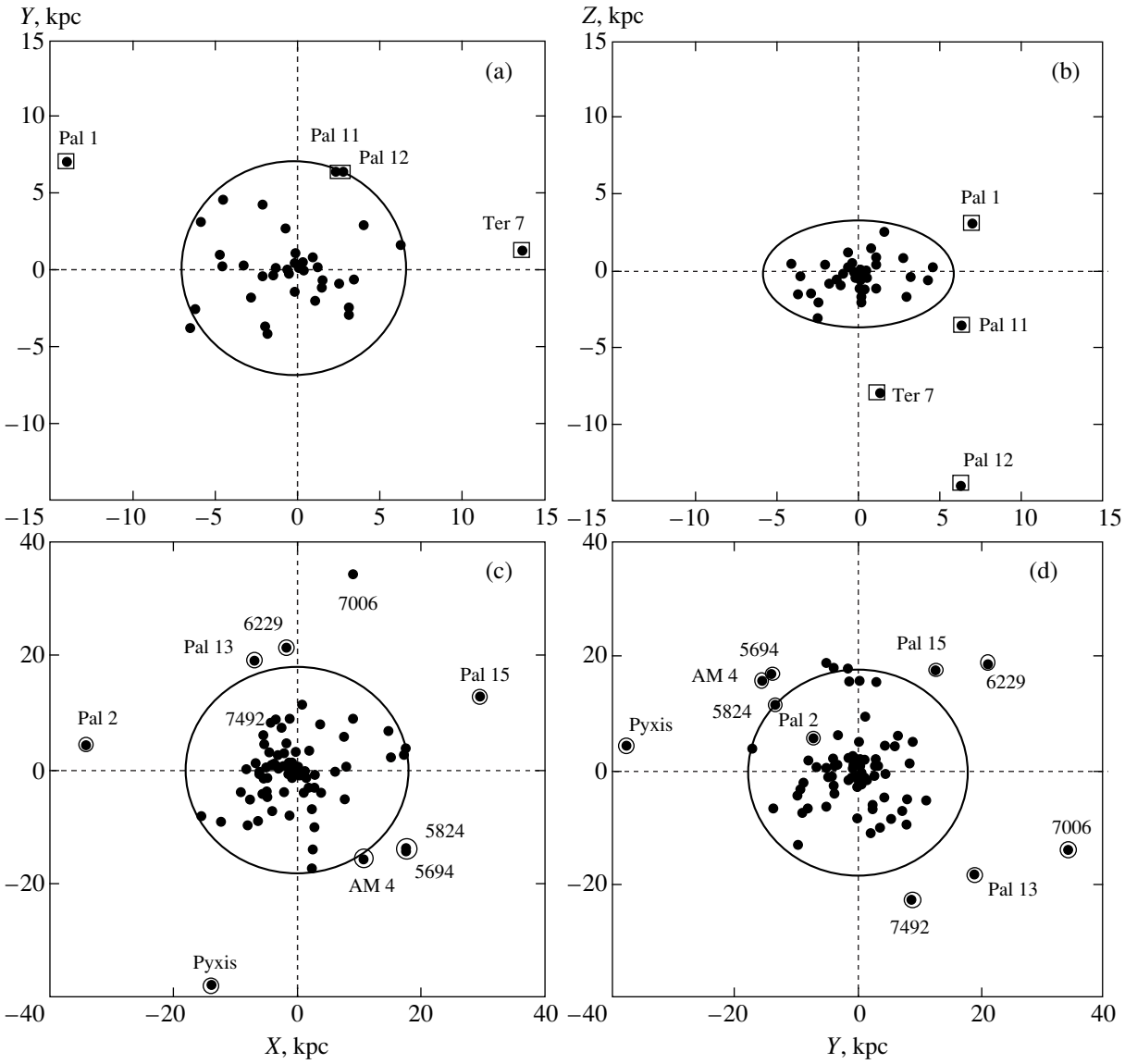


Fig. 2. The distribution of globular clusters projected onto the *XY* [(a) and (c)] and *YZ* [(b) and (d)] planes for metal-rich clusters of the thick disk with [Fe/H] ≥ –1.0 [(a) and (b)] and metal-poor clusters of the halo [(c) and (d)]. The closed curves are upper envelopes drawn by eye. Clusters lying far from the central concentrations in the diagrams are outlined and their numbers indicated.

Properties of GC Groups with Different Metallicities

Figure 3 presents diagrams illustrating the relations between some GC parameters and heavy-element abundance. Abrupt changes in the velocity and distance dispersions, maximum distances of the cluster orbits from the Galactic center and plane (characterizing the total cluster energy), and orbital eccentricities are immediately apparent near the [Fe/H] value separating the halo and disk. This sharp transition suggests a real separation of the GC population into two discrete groups. These groups differ in other parameters as well. For example, all the disk clusters have extremely red HBs, with virtually no normal stars on the HB on the high-temperature side of the gap occupied by variable stars (Fig. 3a). By contrast, the halo clusters can have any HB type (from an extremely red to extremely blue). Age and metallicity are well correlated ($r = -0.5 \pm 0.1$): on average, age decreases with increasing metallicity. Parameters characterizing the internal states of disk and halo clusters do not differ within the errors (Table 2), albeit there are certain systematic effects. We will discuss these in more detail in the next section.

Rotation of Subsystems

The two GC metallicity groups have different angular momenta. If a system is assumed to rotate differentially with constant linear velocity V_{rot}, this velocity can

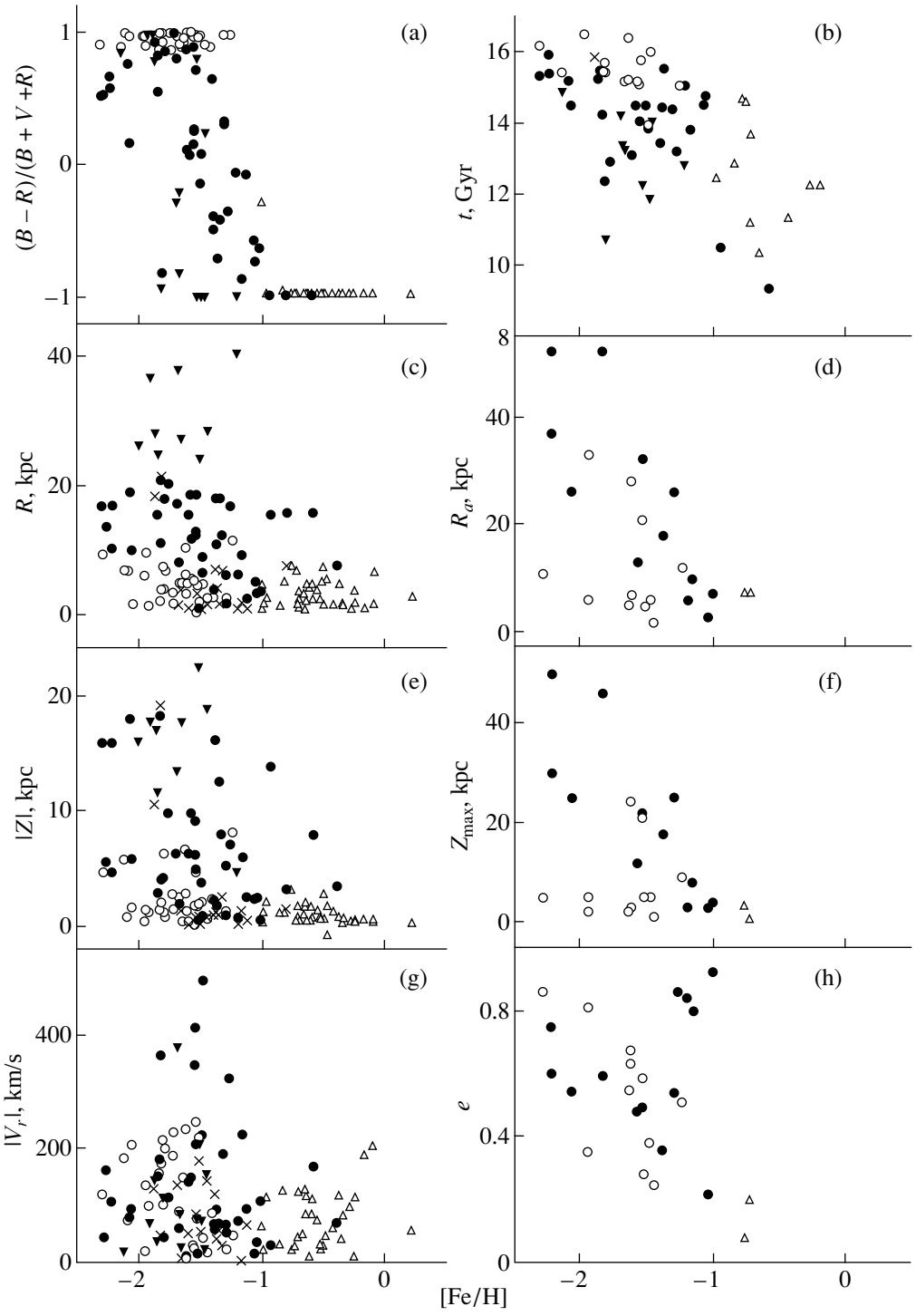


Fig. 3. Relations between the metallicity and other GC parameters. The open triangles are thick-disk clusters; open circles, old halo clusters; filled circles, young-halo clusters; filled triangles, corona clusters; and crosses, metal-poor clusters with unknown $(B-R)/(B+V+R)$. Only clusters with $R < 60$ kpc are shown in (c). A sharp jump near [Fe/H] = –1.0 is seen in all plots.

be derived via least-squares fitting, given only the cluster distances, sky positions, and radial velocities reduced to those for an observer at rest coincident with the Sun (see [2] for details). Figure 4 shows $\cos\psi - V_S$ kinematic diagrams for all Galactic GCs and for the disk and halo subgroups. The straight lines are least-squares regression fits whose slopes yield the rotational velocities of the corresponding subsystems.

The large scatter of data points in the diagrams, which is due to the intrinsic velocity dispersion and

Table 2. Characteristic parameters of GC subsystems

Parameter	Galaxy	Disk	Halo ([Fe/H] < –1.0)	Old halo	Young halo	Corona
$\langle$[Fe/H]$\rangle$	−1.30 ± 0.05(139)	−0.56 ± 0.05(37)	−1.61 ± 0.03(98)	−1.71 ± 0.05(30)	−1.56 ± 0.07(31)	−1.65 ± 0.06(16)
$\sigma_{[Fe/H]}$	0.56 ± 0.03	0.28 ± 0.03	0.30 ± 0.02	0.26 ± 0.03	0.38 ± 0.05	0.26 ± 0.05
$\langle(B-R)/(B+V+R)\rangle$	0.03 ± 0.08(111)	−0.97 ± 0.03(27)	0.42 ± 0.08(80)	0.95 ± 0.01(30)	0.06 ± 0.096(31)	−0.11 ± 0.23(16)
X_0, kpc	4.5 ± 0.5	3.0 ± 0.5	5.5 ± 0.5	3.5 ± 1.0	7.0 ± 3.0	28 ± 1
Y_0, kpc	4.0 ± 0.5	2.0 ± 0.5	4.5 ± 0.5	2.5 ± 0.5	6.5 ± 1.5	23 ± 2
Z_0, kpc	4.5 ± 1.0	1.0 ± 0.2	5.5 ± 1.0	2.5 ± 0.5	8.5 ± 1.5	~30
$\langle\lvert V_r\rvert\rangle$, km/s	107 ± 8(126)	72 ± 9(31)	120 ± 10(92)	113 ± 15(29)	144 ± 23(30)	111 ± 25(14)
$\langle R_a\rangle$, kpc	17 ± 3(25)	7 ± 0(2)	18 ± 3(23)	12 ± 3(11)	24 ± 5(12)	–
$\langle e\rangle$	0.53 ± 0.05(25)	0.13 ± 0.04(2)	0.57 ± 0.04(23)	0.53 ± 0.06(11)	0.59 ± 0.06(12)	–
V_{rot}, km/s	61 ± 25(126)	165 ± 38(31)	23 ± 30(92)	77 ± 33(29)	−23 ± 54(30)	263 ± 221(14)
σ_V, km/s	119 ± 8	88 ± 15	124 ± 17	129 ± 19	140 ± 18	126 ± 24
d([Fe/H]/dR)	−0.025 ± 0.005	0.014 ± 0.022	−0.006 ± 0.003	−0.031 ± 0.016	−0.037 ± 0.010	−0.000 ± 0.002
d([Fe/H]/d$\lvert Z\rvert$)	−0.047 ± 0.009	0.163 ± 0.060	−0.016 ± 0.005	−0.026 ± 0.024	−0.034 ± 0.011	−0.000 ± 0.002
$\langle r_h^*\rangle$, pc	3.3 ± 0.2(138)	2.6 ± 0.3(35)	3.6 ± 0.2(95)	3.4 ± 0.3(30)	4.0 ± 0.4(31)	3.9 ± 0.6(15)
$\langle\log(M/M_\odot)\rangle$	5.2 ± 0.1(144)	5.2 ± 0.1(36)	5.3 ± 0.1(98)	5.4 ± 0.1(30)	5.5 ± 0.1(31)	4.8 ± 0.2(16)
$\langle t\rangle$, Gyr	14.2 ± 0.3(64)	12.5 ± 0.5(12)	14.3 ± 0.2(52)	15.5 ± 0.5(15)	14.1 ± 0.3(26)	13.4 ± 0.4(11)
$\langle C\rangle$	1.6 ± 0.1(139)	1.7 ± 0.1(35)	1.6 ± 0.1(96)	1.6 ± 0.1(30)	1.5 ± 0.1(31)	1.2 ± 0.1(16)
$\langle\log\rho_0\rangle$	3.2 ± 0.2(71)	3.5 ± 0.7(4)	3.2 ± 0.2(66)	3.8 ± 0.3(23)	3.2 ± 0.3(30)	1.6 ± 0.6(10)

Notes: (1) Only clusters with $R < 60$ kpc were used to compute the spatial parameters and metallicity gradients for all subsystems except the corona. (2) The numbers in parentheses give the number of clusters for which the corresponding parameter is known. (3) The X_0 scale length of the old halo is strongly distorted by the observed asymmetry of the distribution of the corresponding clusters projected onto the XY plane (Fig. 4b).

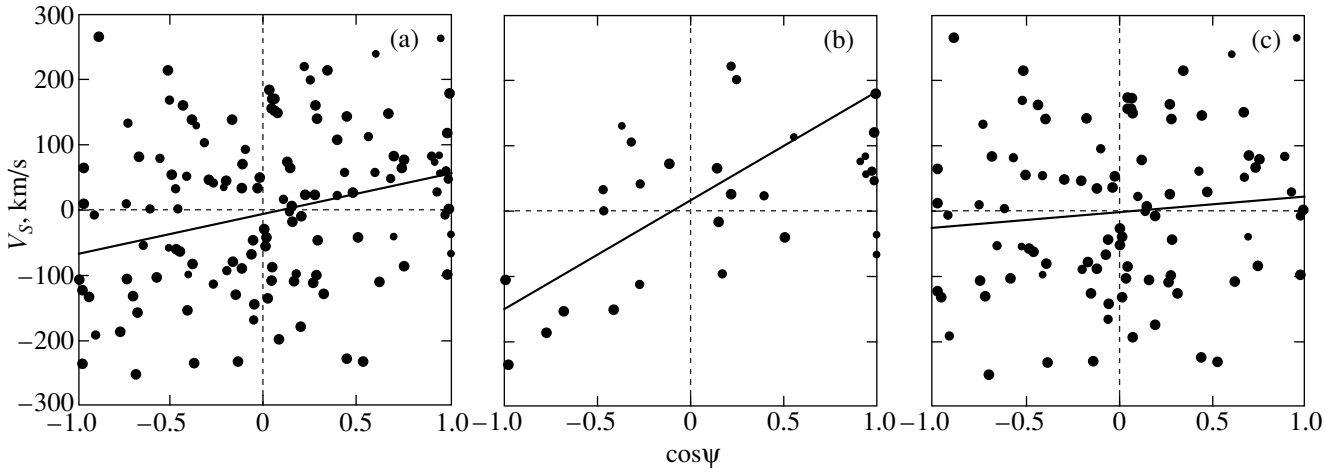


Fig. 4. Kinematic diagrams for (a) all Galactic GCs, (b) thick-disk clusters, and (c) halo clusters. The dot sizes are proportional to the weights assigned. The straight lines are rms regression fits and their slopes yield the rotational velocity of the corresponding subsystem, V_{rot}.

large distance errors, and the small number of objects make the rotational velocities inferred for the GC subsystems quite uncertain. The most difficult case is that of the thick disk, since most of the disk clusters are located near the Galactic center and are thus subject to strong extinction, which distorts the apparent distance moduli. Even a relatively small distance error for a cluster located near the Galactic center translates into a very large error in cos ψ. Therefore, V_{rot} cannot be determined correctly without allowing for the errors in cosψ. The errors in the measured radial velocities and HB magnitudes of the clusters also contribute to the uncertainty in V_{rot}; however, here we allow only for errors in the color excesses $E(B-V)$ since their contribution is dominant in our case. To estimate the errors in the distance moduli, we used the following empirical relation proposed by Harris [30]:

$$\delta(m-M) = 0.1 + 0.3E(B-V).$$

We then transformed the resulting $\delta(m-M)_i$ into distance errors δR_i and computed the variation of cosψ for each cluster by adding and subtracting the error δR_i to and from the cataloged distance. We used the inverses of the mean errors δcosψ as weights. (Unit weights were assigned if $\delta\cos\psi \le 0.05$.) We used these weights when analyzing all the kinematic diagrams in this paper. It follows from Fig. 4 and Table 2 that the resulting difference in the disk and halo rotational velocities exceeds 2σ (the metal-rich subsystem rotates significantly faster). It is also evident from Table 2 that the residual velocity dispersion about the direct regression line, σ_V, for the disk clusters is significantly (i.e., by an amount exceeding the rms error) lower than σ_V for the halo clusters. Douphole *et al.* [26] determined the orbital elements for 25 Galactic GCs, of which 23 belong to the spherical subsystem and only two to the disk. The rotational parameters for the halo subsystem derived in [26] are in good agreement with the corresponding quantities in Table 2 (the rotational velocity $\langle\Theta\rangle = 24 \pm 29$ km/s and $\sigma_\Theta = 137 \pm 20$ km/s). Both disk clusters from [26] have almost the same rotational velocity $\langle\Theta\rangle = 187 \pm 3$ km/s, which also coincides with the corresponding velocity in Table 2 within the errors. Thus, the rotational parameters of GC subsystems with different metallicities differ markedly, and this can naturally explain the shape of each of these subsystems.

Metallicity Gradients

Figure 5 shows *R*–[Fe/H] and |*Z*|–[Fe/H] diagrams for the entire Galactic GC population. From top to bottom, the straight lines are orthogonal regression fits for the GC systems of the disk, entire Galaxy, and halo within $R < 60$ kpc, respectively. The slopes of the regression lines are equal to the corresponding radial and vertical metallicity gradients in Table 2. We can see that the sample of Galactic GCs as a whole demonstrates negative metallicity gradients that are consistent with numerous estimates from other studies. The two correlation coefficients are roughly the same, and are equal to $r \approx -0.40 \pm 0.07$. As noted above, both diagrams exhibit an abrupt change in the spatial characteristics near [Fe/H] ≈ –1.0, indicating that most of the gradients are due to the presence of two GC subsystems with different chemical compositions and spatial distributions. The diagrams for the two GC groups selected according to metallicity, indeed, yield radial gradients that are equal to zero within the errors (Table 2). This result also supports the well-known fact that the radial metallicity gradient is zero beyond the solar circle (where only halo clusters are found). However, the vertical gradients for both groups differ from zero, though the value for the halo group is very small, and much smaller than the gradient for the entire sample. However, the disk subsystem has a very large gradient. Note

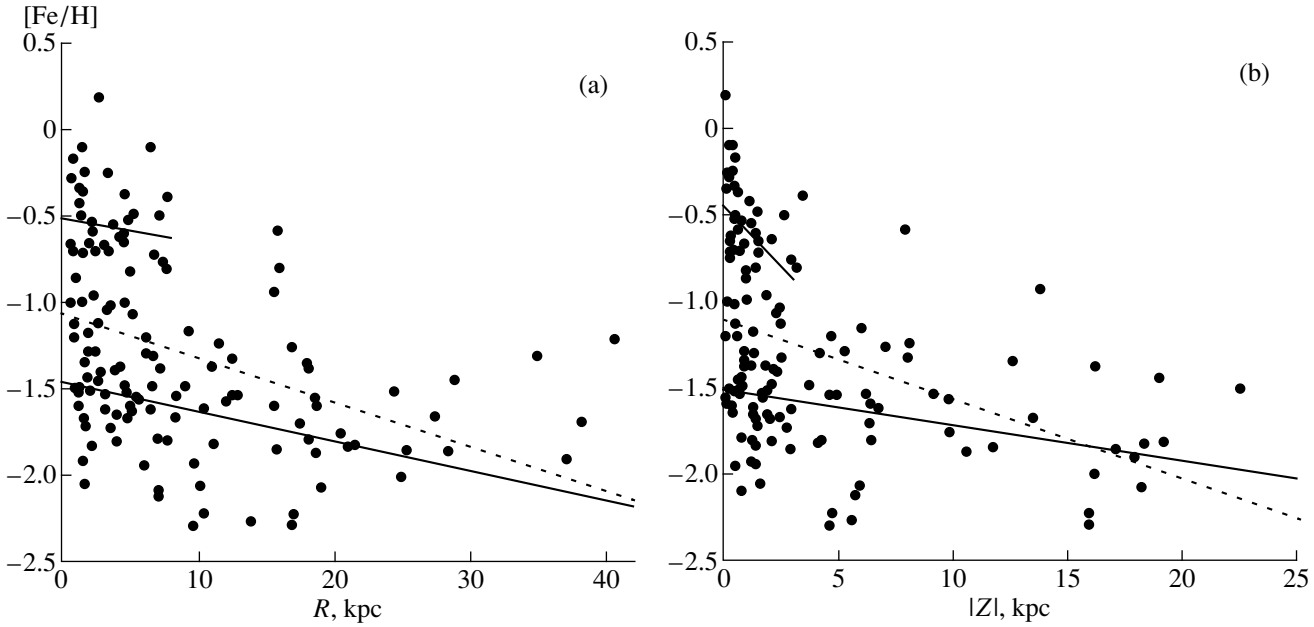


Fig. 5. Relation between metallicity and (a) cluster Galactocentric distance and (b) distance from the Galactic plane. The straight lines are the rms regression fits for the thick-disk, entire Galaxy, and halo from top to bottom, respectively. The slopes of the lines yield the corresponding metallicity gradients.

that the large [Fe/H] gradient for the thick-disk subsystem is not surprising in view of the fact that an even higher value has been derived for the young thin-disk subsystem (see, e.g., [31]).

It thus follows from our catalog that the metallicity-selected GC subsystems have substantially different ages, spatial distributions, rotational velocities, velocity dispersions, and HB morphologies. Let us now consider the complex structure of the halo.

4. GLOBULAR-CLUSTER POPULATIONS IN THE HALO

[Fe/H]–(B – R)/(B + V + R) Diagram

One striking feature of Fig. 3a is the very strong concentration of the color indices of the HBs of halo GCs toward extremely blue values. The distribution of this index shows that the number of GCs increases abruptly at $(B-R)/(B+V+R) \approx 0.85$, which roughly divides the halo population into two subgroups. In his pioneering work, Zinn [3] included clusters located within the solar circle in the old halo population. His sample contains most extremely blue clusters and some GCs located within a narrow band along the upper envelope (in the color parameter) of the [Fe/H]–$(B-R)/(B+V+R)$ diagram. The remaining clusters were assigned to the young halo. In addition, Zinn [3] left a number of very metal-poor clusters with intermediate colors lying well beyond the solar circle in the old-halo group (see Figs. 6c and 6d below). This approach overestimated the size of the old-halo population. Therefore, we consider it more natural to leave only extremely blue clusters in the old-halo population, thereby excluding doubtful members of this subsystem.

Spatial Distribution

Figure 6 shows the spatial distributions of the two groups selected according to HB color. We can see that the old-halo clusters are appreciably concentrated toward the Galactic center, and their distribution is enclosed by a circle of radius ≈9 kpc in the *YZ* plane. In the projection onto the Galactic plane, a circle of the same radius describes only the half of the space nearest to the Sun, whereas the more distant half-space is free of clusters. This asymmetry can hardly be due to excessive extinction near the Galactic center, since disk and young-halo clusters, as well as more distant clusters of the old-halo subsystem, are observed in these locations. On the whole, this problem requires more detailed analysis, and we will not consider it further here.

The squares in Figs. 6a and 6b indicate clusters that lie well beyond the central concentration for this group, but are within the radius of the young-halo population, as shown in Figs. 6c and 6d. The circles in Figs. 6c and 6d indicate clusters located at Galactocentric distances ≥20 kpc. The distribution of young-halo clusters in the *YZ* plane fits well inside a circle of radius $r \approx 19$ kpc. The clusters occupy a much smaller area in the *XY* plane, and their distribution can be described by an elongated figure with semi-axes equal to 18 and 10 kpc (which to some extent can be considered a Galactocentric ellipse). The semi-minor axis is perpendicular to the *Z* axis and makes an angle of about 30° to the *X* axis. Interestingly, the clusters of the Galactic corona have a

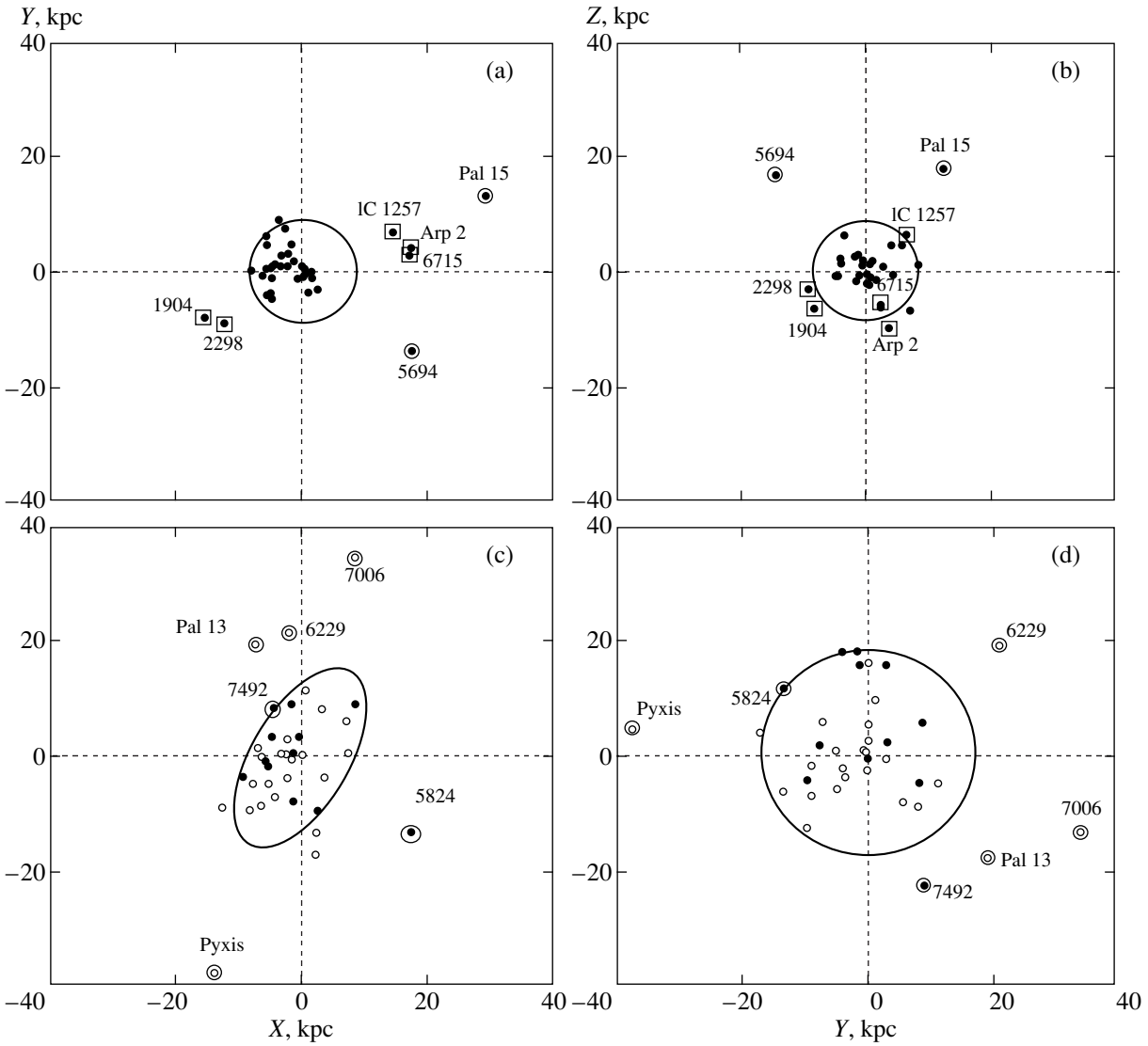


Fig. 6. The distributions of halo clusters projected onto the *XY* [(a) and (c)] and *YZ* [(b) and (d)] planes for old-halo clusters with $(B-R)/(B+V+R) \geq 0.85$ [(a) and (b)] and young-halo clusters [(c) and (d)]. The open and filled circles in (c) and (d) show clusters with $(B-R)/(B+V+R) \leq 0.32$ and $(B-R)/(B+V+R) > 0.32$, respectively. Other notation is the same as in Fig. 2.

somewhat similar distribution. This is partially (within 40 kpc) visible in Figs. 2c, 2d, 6c, and 6d. It also follows from Table 2 that the size of the young-halo subsystem in the *Z* direction substantially exceeds its sizes in the *X* and *Y* directions, which are approximately equal to each other.

Scale Heights of the GC Subsystems

Figure 7 shows $|Z|$ distributions for thick-disk, old-halo, and young-halo clusters. The curves are exponential least-squares fits:

$$n(Z) = Ce^{-Z/Z_0},$$

where Z_0 is the scale height. We similarly computed the scale lengths X_0 and Y_0. The corresponding values for all the GC subsystems are listed in Table 2.

GC Horizontal Branch Colors

Let us return once more to the boundary separating the GC population into the old and young halo subsystems. For comparison, different symbols in Figs. 6c and 6d show young-halo clusters, which we further subdivided into two color groups separated by $(B-R)/(B+V+R) = 0.32$. This is not an arbitrarily boundary: tests showed that it precisely corresponds to an Oosterhoff-type dichotomy in the clusters (red and blue clusters correspond to types I and II, respectively). It is clear from these figures that these intermediate-color subgroups have similar spatial distributions, and apparently both belong to the young halo. To test this conclusion and refine the position of the boundary separating the old and young halos, we analyzed the dependence of the HB color on other cluster parameters. The corre-

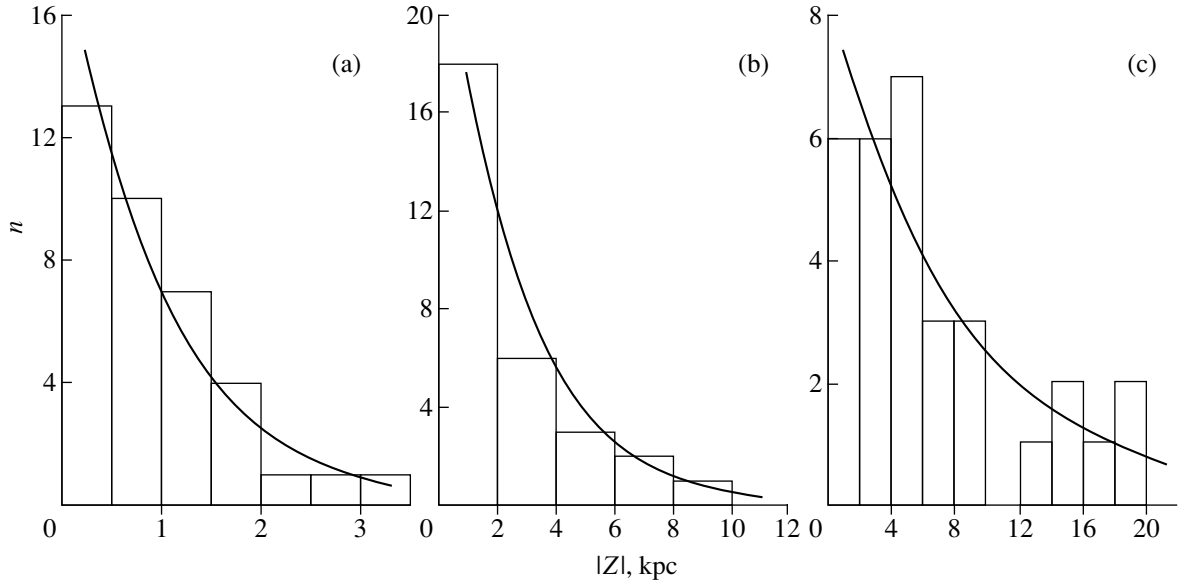


Fig. 7. Distributions of height above the Galactic plane for (a) thick-disk, (b) old-halo, and (c) young-halo clusters. The curves are exponential least-squares approximations of the corresponding distributions.

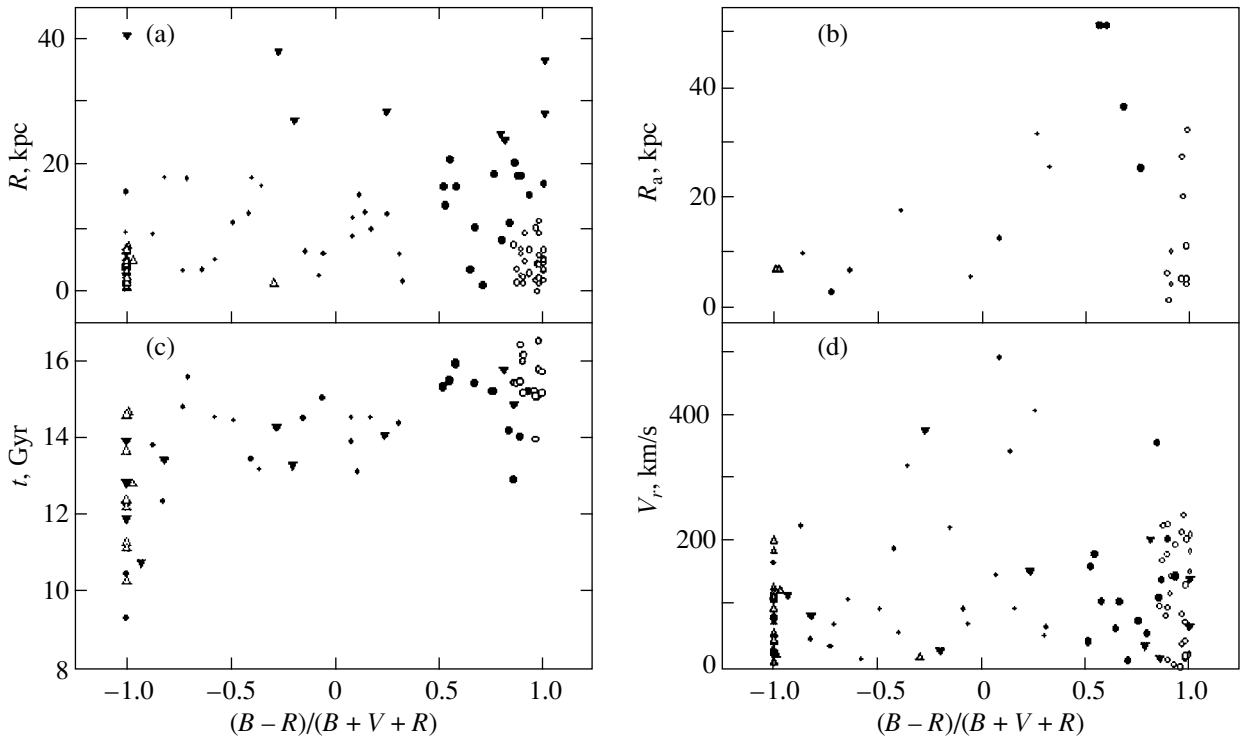


Fig. 8. Relations between GC parameters and the horizontal-branch color index. The open circles correspond to the old halo, filled circles to the young halo (large filled circles correspond to clusters with $(B-R)/(B+V+R) > 0.32$), and filled triangles to the corona.

sponding diagrams are shown in Fig. 8. (Here, the five distant GCs shown by squares in Figs. 6a and 6b are included in the young-halo population.)

Comparison of the positions of the points in the first diagram demonstrates that the degree of concentration toward the Galactic center does change abruptly at the value $(B-R)/(B+V+R) = 0.85$ separating the young and old halo populations. Moreover, some clusters with $(B-R)/(B+V+R) > 0.32$ in Fig. 8b have apogalactic orbital radii appreciably exceeding those for redder

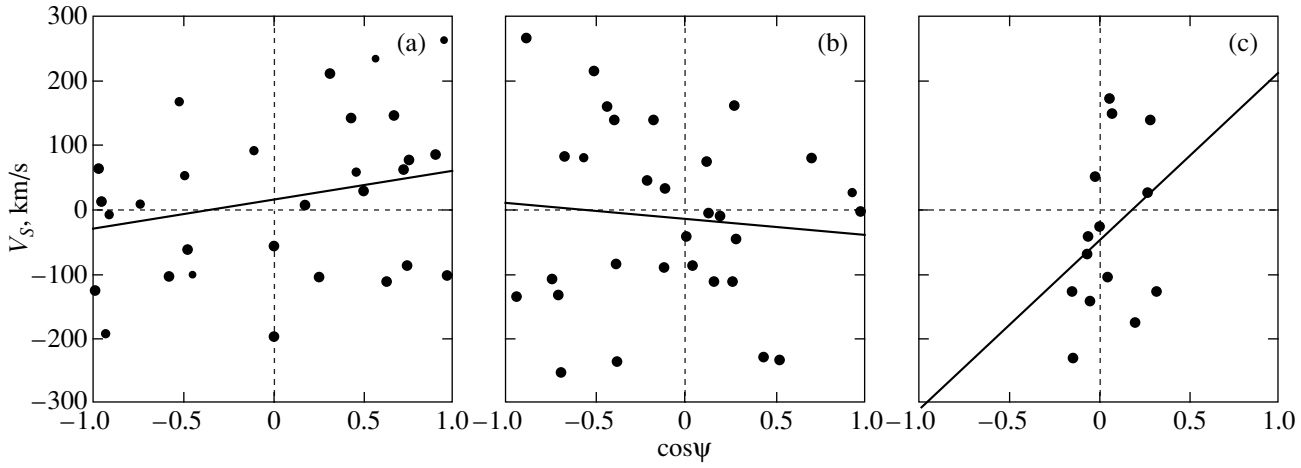


Fig. 9. Kinematic diagrams for (a) old-halo, (b) young-halo, and (c) corona clusters. Notation is the same as in Fig. 4.

clusters (and even exceeding the size of the young-halo subsystem). It is evident from Fig. 8d that the velocity dispersion for this group is also higher than for the old halo. [Note that, logically, clusters with $R_a > 20$ kpc should be assigned to the corona (see below); however, we have not done this because of the high probability of substantial errors in the orbital elements.] In addition, the clusters of this subgroup, whose average age is 15.3 Gyr (virtually the same as that of the old halo in Table 2), have a very low age dispersion of only $\sigma_t = 0.5 \pm 0.2$ Gyr (Fig. 8c). The fact that these clusters also have approximately the same metallicities ([Fe/H] = –1.93 ± 0.10, see Fig. 3a) and eccentricities ($e = 0.62 \pm 0.05$) suggests that most are members of a homogeneous group, and must have been captured simultaneously by the Galaxy.

The figures also show that the greatest distances (exceeding the size of the old halo), smallest ages (younger than that of the young halo), and highest velocities (higher than those for the old halo) are found for clusters located near the boundary $(B - R)/(B + V + R) = 0.85$. As a result, we assigned all clusters with intermediate-type HBs and distant clusters ($R > 10$ kpc) with extreme HB colors to the young-halo subsystem.

For now, we consider the most distant clusters ($R \geq 20$ kpc) to be a separate group, which we tentatively call the Galactic "corona". We can see from Table 2 that the corona clusters differ significantly from the clusters of all the other subsystems in both the spatial volume they occupy and their physical parameters. We will attempt to interpret these differences below.

Rotation of the Halo Subsystems

The kinematic diagrams in Fig. 9 show that all the halo subsystems differ dramatically in their rotational velocities. The parameters V_{rot} and σ_V are given in Table 2. The old halo exhibits a fairly well-defined (at the 1.7σ level) prograde rotation whose parameters agree reasonably well with those derived by other authors (see, e.g., [3], where $V_{rot} = 74 \pm 39$ km/s and $\sigma_V = 129 \pm 19$ km/s). Such a low rotational velocity for this subsystem (almost half that for the disk) is consistent with its spherical shape.

The diagram in Fig. 9b indicates retrograde rotation of the young halo and a rather high velocity dispersion (Table 2). The substantial deviation of the regression line from the coordinate origin suggests inhomogeneity of this system. It appears that the rotation of the system about the Z axis must be not entirely axially symmetric. The flattening of the young halo along an axis that does not coincide with the Z axis could only come about as the result of a chance coincidence of the orbital inclinations of clusters whose origin is not related to that of the protogalaxy. The most plausible hypothesis seems to be that these clusters have an extragalactic origin, and were captured one after another as the Galaxy (with some definite orientation) crossed regions of cluster concentration. The enormous prograde rotational velocity of the corona, together with its relatively low σ_V and the large deviation of the line in Fig. 8c from the origin, also testify to an extragalactic origin for the corona clusters. The prograde rotation of this group is beyond question, since it follows from Table 2 that V_{rot} differs from zero by more than σ.

Metallicity Gradients

The different symbols in Fig. 10 indicate clusters of the old and young halos. The solid and dashed straight lines are rms regression fits based on the data for GCs of the young and old halo, respectively. The corresponding metallicity gradients are given in Table 2. For the old halo, we computed both gradients excluding NGC 288 (indicated by a square in all diagrams), which is both the most metal-rich and the most distant cluster in the group. It follows from Fig. 10 and Table 2 that the old-halo subsystems all have the same values of corresponding gradients. Moreover, the gradients in the two

directions coincide within the errors in all the subsystems. Our result for the young halo is at variance with Zinn's [3] conclusion that this GC group exhibits no radial metallicity gradient. The discrepancy is due to our different cluster subsystem selection criteria: Zinn [3] included neither the group of distant ($R > 10$ kpc), metal-poor clusters with HB color parameters $0.32 \le (B-R)/(B+V+R) \le 1$ nor the group of the most metal-rich halo clusters with $-1 < (B-R)/(B+V+R) < 0.32$, which (being the extreme points in the diagrams) determine to a considerable degree the gradients for this subsystem. Our results thus show that, at all Galactocentric distances, the young-halo clusters are, on average, ~0.3 dex richer in metals than old-halo clusters. The most extended group—the corona—exhibits no gradients in either direction.

Physical Parameters

It is immediately apparent from an analysis of Table 2 that all the mean physical parameters of the corona clusters differ radically from those for the other groups. This group is the most distant, and therefore we should check whether the parameters exhibit any Galactocentric trends. If they do, this large difference in parameters could simply reflect changes of the internal structure of clusters with Galactocentric distance. Figure 11 shows relationships between some physical parameters and Galactocentric distance. Van den Bergh [32] showed that, on average, the cluster sizes increase with Galactocentric distance, and obey the relation $r_h \propto \sqrt{R}$ within 40 kpc. The new distance scale and refined angular sizes of GCs confirm that this relation remains valid even for the most distant clusters. The direct rms regression line in $(\log R, \log r_h)$ in Fig. 11 has a slope of 0.44 ± 0.04, in satisfactory agreement with the relation given by van den Bergh [32]. The approximately constant scatter along the entire regression line indicates that the cluster sizes reduced to the solar Galactocentric distance with van den Bergh's [32] formula $r^*_{hi} = r_{hi}(R_i/R_G)^{-0.5}$ are roughly the same for all subsystems. In other words, the large observed radii of the corona clusters are in no way indicative of a different origin.

It follows from Table 2 that, on average, the corona clusters have the smallest masses. We can see from Fig. 11b that the GC masses indeed decrease as the distance increases from 10–120 kpc. However, the opposite relation is found for nearby clusters: the cluster masses slightly increase with Galactocentric distance. The large circles in the diagrams show the mean masses for narrow R intervals, making this pattern immediately apparent. The relationship becomes even more conspicuous if we recall that $R = 10$ kpc is near the boundary of the old halo, and analyze the mass–distance relations separately for the disk and old-halo groups (open symbols in the diagram), on the one hand, and for the young-halo and corona groups (filled symbols), on the other hand.

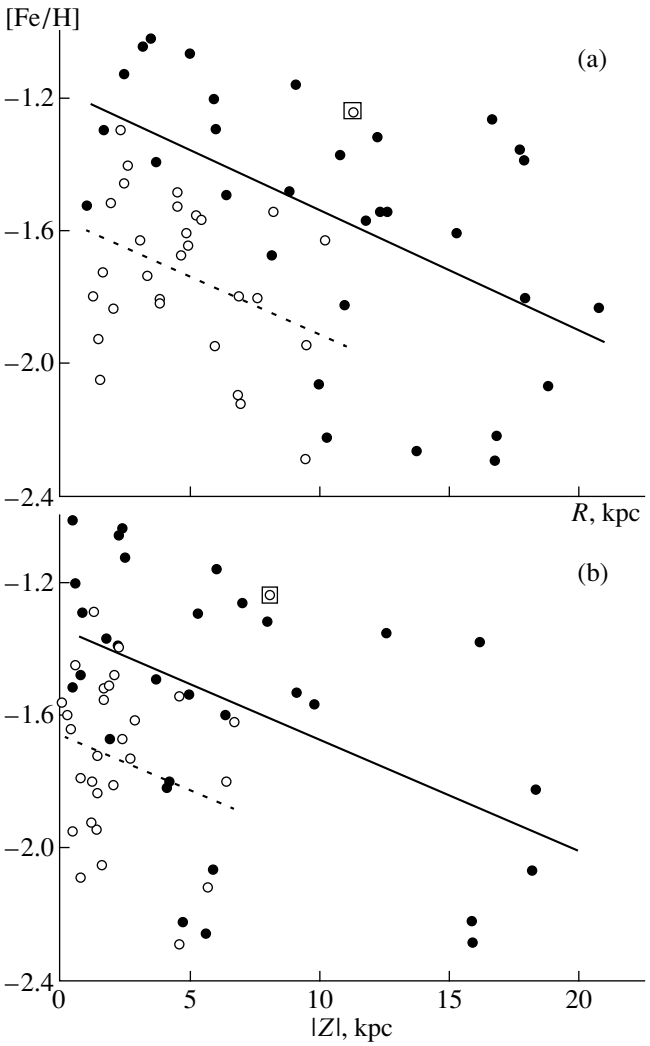


Fig. 10. Dependence of cluster metallicity on (a) Galactocentric distance and (b) distance from the Galactic plane. The filled and open circles correspond to the young and old halos, respectively, and the straight lines are the corresponding rms regressions. The small square shows the old-halo cluster excluded from computations based on a 3σ criterion.

The fact that the mean mass of the more distant old-halo clusters in Table 2 exceeds that of clusters belonging to the more compact thick-disk subsystem is enough for us to conclude that the selected groups show different trends. The first two groups, whose clusters are genetically associated with the Galaxy, exhibit a small, positive, radial mass gradient, whereas the GCs in the two other groups, which purportedly have an extragalactic origin, exhibit a very strong, negative, radial mass gradient (see the direct rms regression lines in the diagram).

Table 2 indicates that the corona GCs have the lowest densities. These reflect the overall trend in the R–$\log\rho$ diagram, with a negative gradient for the internal clus-

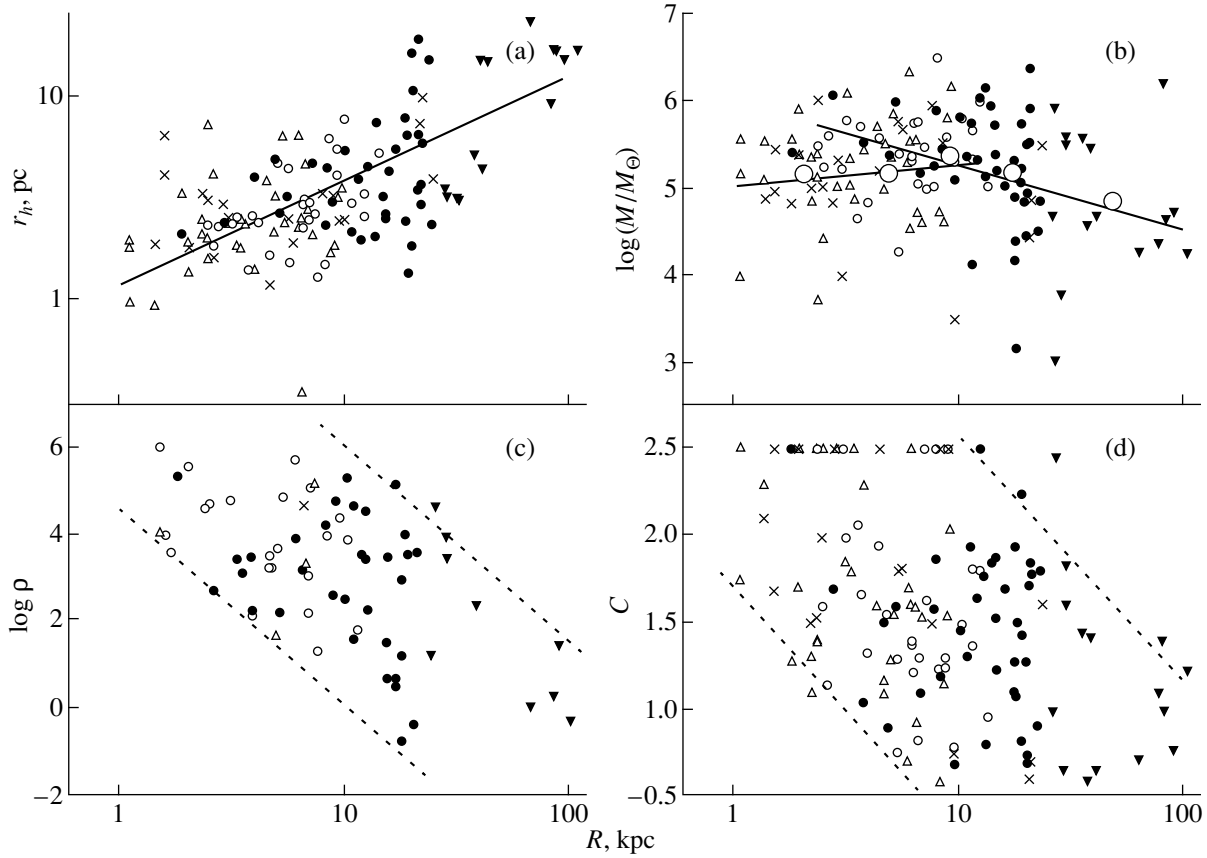


Fig. 11. Galactocentric distance dependences of the physical parameters of globular clusters. Notation is the same as in Fig. 3. The straight line in (a) is the rms regression. The large circles in (b) show mean masses within narrow R intervals, and the solid lines show the rms regression fits for thick-disk and old-halo (left) and young-halo and corona (right) clusters. The dashed lines in (c) and (d) are upper and lower envelopes drawn by eye.

ter density. In this sense, the corona clusters are linked to the other clusters through a common relation (the dashed lines in Fig. 11c are upper and lower envelopes drawn by eye). According to Table 2 and Fig. 11d, the central concentration also monotonically decreases with Galactocentric distance. Thus, of the four physical parameters considered, only the cluster mass is independent of distance and exhibits somewhat different Galactocentric trends in the subsystems lying inside and outside the solar radius.

Cluster Ages

It follows from Figs. 3b and 8c that the cluster ages are strongly correlated with two main parameters—metallicity and HB-morphology index. It is therefore of interest to analyze how other parameters depend on age. The corresponding diagrams are shown in Fig. 12. Three diagrams have nonzero correlation coefficients. In this figure, we reduced the cluster radii to the solar Galactocentric radius in accordance with the relation of [32], thereby removing the Galactocentric distance trend in cluster size. Figure 12a shows that the reduced cluster radii exhibit no age dependence. The radii of clusters of all ages are confined to the interval $1 < r_h^* < 6$ pc. Only a few clusters of the young halo and corona populations have larger sizes. In the $t - \log(M/M_\odot)$ diagram, the cluster masses clearly grow with age, with a correlation coefficient of $r = 0.4 \pm 0.1$. This may suggest that old clusters live longer; however, this is inconsistent with the complete lack of young, massive clusters.

Likewise, it is evident from Figs. 12c, 12d that old clusters have, on average, higher densities and more prominent central cores (the correlation coefficients are 0.5 ± 0.1 and 0.2 ± 0.1, respectively). No young clusters with similar parameters are known. At the same time, no objects with low mass, density, and central concentration can be found among the oldest clusters. The same pattern is demonstrated by corresponding parameters in all the GC subsystems. The parameters of individual clusters of a given age exhibit a rather large scatter, whatever subsystem they belong to.

It appears that, of all the parameters considered, mass is the only one that remains constant from the time of cluster formation and is independent of the position of the cluster in the Galaxy. If we suppose that some clusters are genetically unrelated to the Galaxy and have been cap-

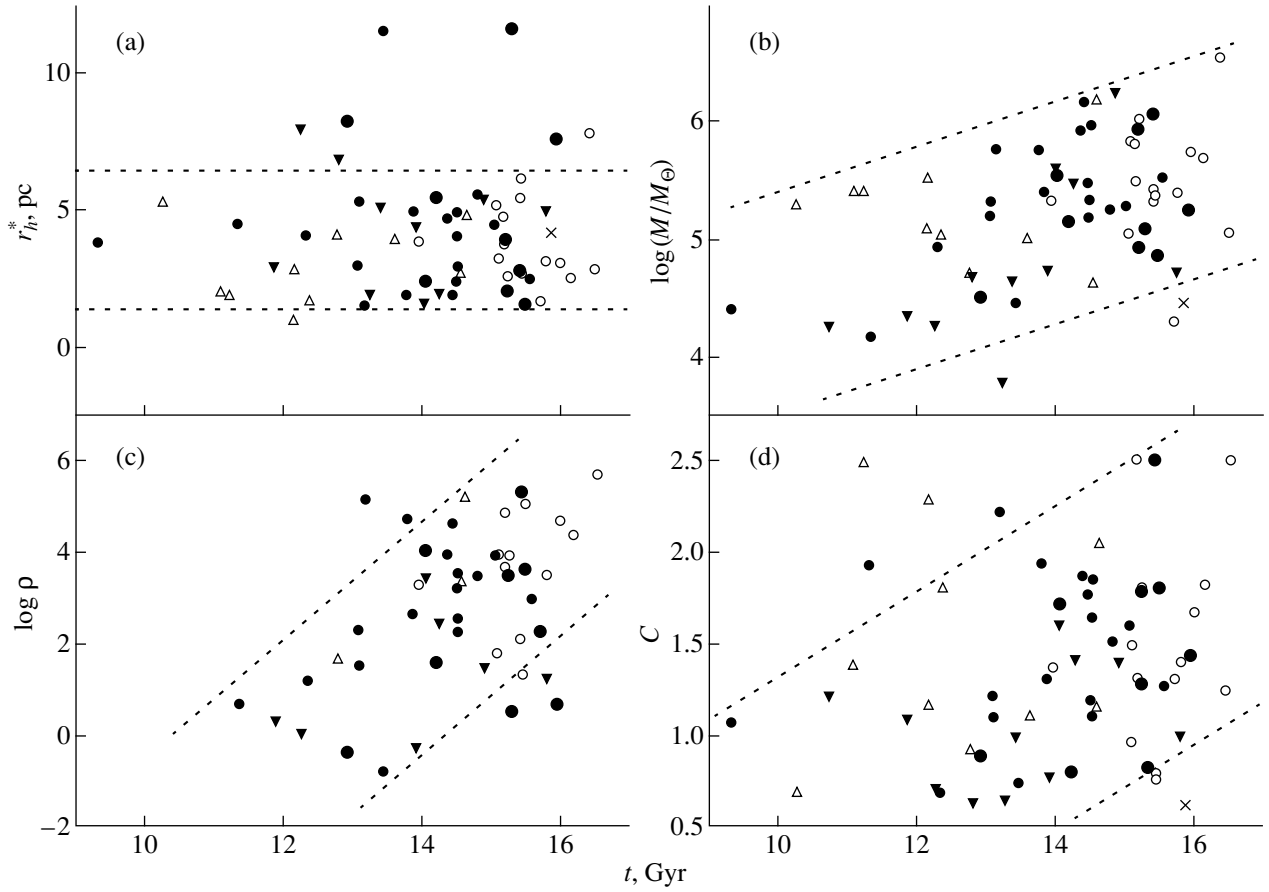


Fig. 12. Age dependences of the physical parameters of globular clusters. Notation is the same as in Fig. 11. The dashed lines are envelopes drawn by eye.

tured, we must conclude that, throughout the entire Local Group of galaxies, the most massive GCs formed first, and younger clusters formed later from gas and dust clouds of ever decreasing mass.

Cluster Classification

Let us briefly summarize the criteria used to assign GCs to particular subsystems. The thick-disk subsystem is made up of metal-rich clusters with [Fe/H] ≥ –1.0, marked "1" in Table 1. The four metal-rich GCs lying outside the volume occupied by disk clusters probably belong to the young halo (they are marked "3(1)" in Table 1)[1] Extremely blue, metal-rich clusters with $(B-R)/(B+V+R) \geq 0.85$ form the old halo ("2"). Extremely blue GCs located outside the spherical volume occupied by the old-halo subsystem are assumed to belong to the young halo ("3(2)"). All metal-poor clusters with red and intermediate-color HBs make up the young halo ("3"). GCs of unknown metallicity (or HB color) are either unclassified or tentatively assigned to some subsystem according to the parameter in parentheses.

The GC classification into subsystems proposed here must be considered only tentative. We tried to identify primarily pure examples of members of the thick-disk and old-halo populations in order to reliably determine the characteristic parameters of these subsystems (and vice versa). Burkert and Smith [33] showed that thick-disk GCs do not form a homogeneous group, and can be divided into three discrete subgroups according to their mass and kinematics. Each of the subsystems identified here (especially, the young halo) appears to have a fairly complex structure and requires further study.

5. CONCLUSIONS. A SCENARIO FOR GALACTIC EVOLUTION

Thus, analysis of astrophysical data from recent accurate observations confirms earlier conclusions that the Galactic globular-cluster population is inhomogeneous. The thick-disk and old-halo subsystems identified here based on their heavy-element abundances and HB colors also differ markedly in the size and shape of the spatial volume they occupy, their kinematics, and

[1] The number of the subsystem to which the cluster would have been assigned according to the basic criterion (in our case, metallicity) is given in parentheses.

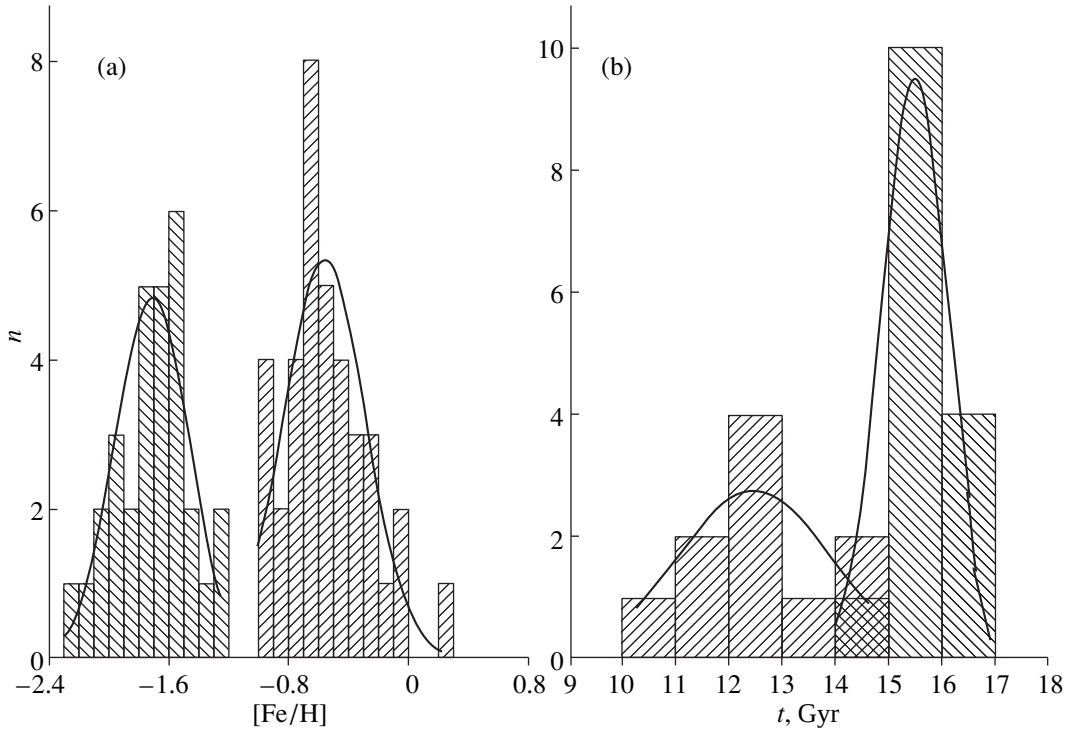


Fig. 13. The (a) metallicity and (b) age distributions of the globular clusters of Galactic origin. Different shadings are used to show the thick-disk and old-halo clusters. The curves are Gaussian fits to the corresponding histograms.

their orbital elements. The combined metallicity function of the two subsystems is a superposition of two more or less equal normal distributions separated by a gap Δ[Fe/H] ≈ 0.2 (Fig. 13a). The age distribution of the same clusters also forms two virtually non-overlapping Gaussians (Fig. 13b).

It follows from Fig. 13b that the age dispersion for the halo clusters is nearly equal to the error in the ages themselves, $(\sigma_t)_{\rm halo} = 0.8 \pm 0.2$ Gyr. The disk-cluster ages exhibit a much larger scatter, $(\sigma_t)_{\rm disk} = 1.4 \pm 0.3$ Gyr. In other words, the halo clusters formed almost simultaneously, whereas the corresponding processes in the disk required at least several billion years. The age gap between subsystems, if any, might be swamped by the errors in the estimated ages.

Figure 14 shows the age dependences of the metallicity, Galactocentric distance, and distance from the Galactic plane for the two subsystems considered. The straight lines in these plots are individual rms regression fits for each of the subsystems and for the Galactic GC sample as a whole. The correlation coefficients indicate that at least the first two dependences for the entire sample are real ($r = 0.8 \pm 0.1$, 0.6 ± 0.1, and 0.3 ± 0.2, respectively). Separately, the thick-disk subsystem exhibits no age–metallicity correlation ($r = 0.3 \pm 0.3$), whereas the old-halo population, in spite of its narrow age spread, shows a weak dependence ($r = 0.4 \pm 0.2$) with a slope that is close to that for the entire sample. Galactocentric distances are completely uncorrelated with age in the old halo ($r = 0.1 \pm 0.3$). However, the corresponding dependence in the disk proved to be rather significant ($r = 0.7 \pm 0.2$) and in agreement with that for the Galaxy as a whole (Fig. 14b). An interesting pattern can be seen in Fig. 14c, which shows an abrupt jump in $|Z|$ between the subsystems rather than an age dependence for cluster height above the Galactic plane, with virtually no correlation within each subsystem.

Analysis of these patterns suggests the following scenario for the early evolution of our Galaxy. The first globular clusters formed when the protogalactic cloud had already collapsed to the size of the modern Milky Way, $R_{GC} \approx 12$ kpc. (It remains unclear, however, whether the slightly reddened, metal-poorest old GCs at distances $R = 15$–25 kpc belong to the old halo.) The old-halo subsystem formed over a short time interval, making it impossible to confidently identify age trends in either cluster size or metallicity. The radial and vertical metallicity gradients in this oldest known Galactic subsystem provide evidence that the first heavy-element enrichment of the gas–dust medium took place before the clusters of this subsystem formed, and that this enrichment was more active near the Galactic center. (Note that the clusters maintain the orbits of their parent protoclouds. Evidence for the absence of relaxation is provided, in particular, by the existence of a large number of fairly old and simultaneously very distant clusters in the young halo.)

Both the spatial and chemical characteristics change abruptly as we move to the thick-disk subsystem. The formation of stars in these two subsystems was appar-

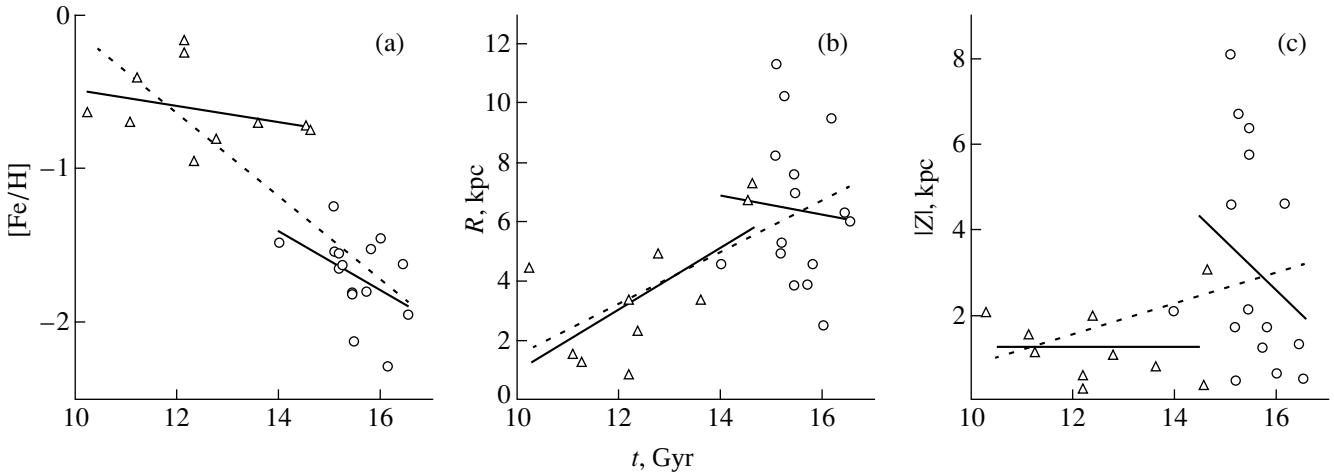


Fig. 14. Age dependences of (a) metallicity, (b) Galactocentric distance, and (c) distance from the Galactic plane for the globular clusters of Galactic origin. Notation is the same as in Fig. 3. The dashed lines are rms regression fits for both subsystems taken together, and the solid lines are the individual regression fits for each subsystem. The squares indicate clusters excluded from the regression solutions according to a 3σ criterion.

ently separated by a substantial time lag, which shows up only as a gap in the age distributions. This time lag enabled the gas and dust clouds to become appreciably enriched in heavy elements (which had time to mix) and to collapse to much smaller sizes before the new generation of globular clusters began to form. As a result, a rather flat, metal-rich, thick-disk subsystem formed. This subsystem acquired a much higher rotational velocity and much lower velocity dispersion compared to the old halo. Neither the mean metallicity nor the thickness of the thick disk changed with time. Star formation in the protoclouds of the clusters of this subsystem began at the largest Galactocentric distances, then began to propagate toward the center. The appreciable collapse of the protodisk cloud after the formation of the halo subsystem resulted in an increase of the rotational velocity and a rapid flattening of the future subsystem. Apparently, the interstellar medium did not undergo deep mixing in the process, since we observe a strong, negative vertical metallicity gradient with virtually no radial gradient in this subsystem.

In other words, our analysis suggests that there was a rapid collapse of the gas–dust medium toward the Galactic plane with simultaneous heavy-element enrichment and partial homogenization, which took place mainly during the period between the formation of the old-halo and thick-disk subsystems. This scenario requires further refinement, since the quality of current observational data is insufficient to enable a deeper understanding of the processes within each subsystem of Galactic globular clusters during its formation.

The young-halo and corona subgroups are made up of clusters (or proloclusters) captured by the Galaxy from the intergalactic medium in various stages of its evolution. They differ from subsystems genetically associated with the Galaxy primarily in the volume they occupy. For example, the sizes of the young halo and the corona are factors of two and ten larger than that of the old halo. Note also the very large uncertainty in our rotational velocities of the subsystems and the great radial-velocity dispersions. The separation of these clusters into two subgroups is somewhat arbitrary, and we have done this only to demonstrate the extreme inhomogeneity of the spatial–kinematic structure of the member clusters. It would be more correct to represent the two subgroups as a single subsystem consisting of clusters randomly captured by the Galaxy, which could be called the "outer halo."

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 00-02-17689).

REFERENCES

1. V. A. Marsakov and A. A. Suchkov, Astron. Zh. **54**, 1232 (1977) [Sov. Astron. **21**, 700 (1977)].
2. R. Zinn, Astrophys. J. **293**, 424 (1985).
3. R. Zinn, in *The Globular Cluster-Galaxy Connection*, Ed. by H. Smith and J. Brodee, ASP Conf. Ser. **48**, 38 (1993).
4. G. S. Da Costa and T. F. Armandroff, Astron. J. **109**, 2533 (1995).
5. W. E. Harris, Astron. J. **112**, 1487 (1996).
6. P. Thomas, Mon. Not. R. Astron. Soc. **238**, 1319 (1989).
7. R. Zinn and M. J. West, Astrophys. J., Suppl. Ser. **55**, 45 (1984).
8. S. van den Bergh, Mem. Soc. Astron. Ital. **54**, 199 (1983).
9. R. F. Webbink, in *IAU Symposium 113: Dynamics of Star Clusters*, Ed. by J. Goodman and P. Hut, p. 541 (1985).

10. E. Carretta and R. G. Gratton, Astron. Astrophys., Suppl. Ser. **121**, 95 (1997).
11. A. V. Mironov and N. N. Samus', Perem. Zvezdy **19**, 337 (1974).
12. I. N. Reid, Astron. J. **114**, 161 (1997).
13. B. Hauck and M. Mermilliod, Astron. Astrophys., Suppl. Ser. **129**, 431 (1998).
14. R. Buonnano, C. E. Corsi, and F. Fusi Pecci, Astron. Astrophys. **216**, 80 (1989).
15. R. G. Gratton, Astron. Astrophys. **147**, 169 (1985).
16. R. Buonnano, C. E. Corsi, and F. Fusi Pecci, Astron. Astrophys. **216**, 80 (1989).
17. A. Sarajedini and G. R. King, Astron. J. **98**, 1624 (1989).
18. B. Chaboyer, A. Sarajedini, and P. Demarque, Astrophys. J. **394**, 515 (1992).
19. A. Sarajedini, Y.-W. Lee, and D.-H. Lee, Astrophys. J. **450**, 712 (1995).
20. B. Chaboyer and Y.-C. Kim, Astrophys. J. **454**, 767 (1995).
21. B. Chaboyer, P. Demarque, and A. Sarajedini, Astrophys. J. **459**, 558 (1996).
22. H. B. Richer, W. E. Harris, C. Fahlman, *et al.*, Astrophys. J. **463**, 602 (1996).
23. M. Salaris and A. Weiss, Astron. Astrophys. **327**, 107 (1997).
24. R. Buonanno, C. Corsi, M. Bellazzini, *et al.*, Astron. J. **113**, 706 (1997).
25. F. Fusi Pecci, F. R. Ferraro, M. Bellazini, *et al.*, Astron. J. **105**, 1145 (1993).
26. B. Douphole, M. Geffert, J. Colin, *et al.*, Astron. Astrophys. **313**, 119 (1996).
27. S. van den Bergh, Astron. J. **110**, 1171 (1995).
28. D. F. Chernoff and S. Djorgovski, Astrophys. J. **339**, 904 (1989).
29. T. E. Armandroff, Astron. J. **97**, 375 (1989).
30. W. E. Harris, personal communication.
31. Yu. G. Shevelev and V. A. Marsakov, Astron. Zh. **72**, 321 (1995) [Astron. Rep. **39**, 284 (1995)].
32. S. van den Bergh, Astrophys. J. **375**, 594 (1991).
33. A. Burkert and G. H. Smith, Astrophys. J. Lett. **474**, L15 (1997).

Translated by A. Dambis

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 685–688. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 772–776.

Parameters for the Dust Envelope of the Protoplanetary Nebula IRAS 18062 + 2410

M. B. Bogdanov

Chernyshevsky State University, ul. Dvadtsatiletiya VLKSM 112a, Saratov, 410071 Russia

Received September 24, 1999

Abstract—We have calculated a model for the dust envelope of the protoplanetary nebula IRAS 18062 + 2410 using observations from the ultraviolet to the far infrared. We assume that the envelope is spherically symmetrical and consists of identical silicate grains with a radius of 0.10 micron, and with the number density of the grains inversely proportional to the square of the distance. The optical depth of the envelope, whose inner boundary is 1.40×10^{-3} pc from the center of the star, is 0.050 at 10 microns. At the inner envelope boundary, the temperature of the dust grains is 410 K and their density is 2.7×10^{-7} cm^{-3}. Using calculations of stellar evolution at the stage following the exit from the asymptotic giant branch, we estimate the dust envelope's expansion velocity to be 12 km/s. The mass-loss rate of the star preceding the ejection of the envelope was about $4.5 \times 10^{-6}\ M_\odot$/yr. The observed excess of far-IR flux is not associated with the continuum radiation of the nebula, and may provide evidence for the presence of dust ejected by the star in earlier stages of its evolution.

1. INTRODUCTION

The interesting object IRAS 18062 + 2410 (= SAO 85766 = HDE 341617 = V886 Her) is associated with a star at the stage of evolution following the asymptotic giant branch and preceding the formation of a planetary nebula [1]. This stage is so short that the evolutionary variation of the effective temperature of the star is detectable during the lifespan of one generation of observers. Studies of such objects can help clarify the origin of the intense stellar wind that leads to the ejection of the stellar envelope, which is important for our understanding the final stages of the evolution of stars of moderate mass and the formation of the interstellar medium.

IRAS 18062 + 2410 has been actively studied with both IRAS and ground-based detectors. Photometric observations at near and middle IR wavelengths revealed an emission feature near 10 micron due to the presence of silicate dust [2]. These data and the far-IR flux suggest that the dust temperature in the circumstellar envelope is about 200 K [3]. In [4], *JHK* photometric data are also presented. The most detailed study of the object was based on *UBV* photometry and spectra obtained with the Russian 6-m telescope [5]. The parameters of the nebula and central star were estimated in [6]; some secular variations of the spectral type and effective temperature were also revealed, and shown to be consistent with the results of evolutionary calculations.

In our study, we calculate a model for the dust envelope of IRAS 18062 + 2410 and estimate the mass-loss rate at the beginning of the expansion of the stellar envelope, which should result in the formation of a planetary nebula.

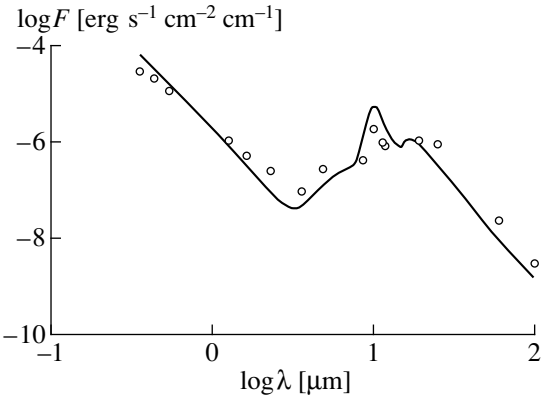
2. DUST ENVELOPE MODEL CALCULATION

Modeling a stellar dust envelope based only on observed flux data is a rather complicated problem due to the large number of unknown parameters describing the central star, dust grains, and envelope geometry. It is natural for such studies to minimize this uncertainty by using *a priori* information from either independent observations or theoretical considerations. Model calculations for envelopes consisting of silicate grains can be found in [1, 7–9]. When calculating the model for the envelope of IRAS 18062 + 2410, we used the parameters of the star presented in [5] and widespread assumptions about the characteristics of the dust grains and geometry of circumstellar envelopes, together with the observed flux data.

We adopted the average *UBV* magnitudes from [5] and the IR photometric data from [2–4]. For the interstellar absorption A_V, we adopted the value $0\overset{m}{.}30$ [5], which is confirmed by the distribution of the color excess $E(B–V)$ in [10]. Our calculation of the absorption in other bands was based on the data from [11]. After taking interstellar absorption into account, we determined the flux $F(\lambda)$ (in erg s^{-1} cm^{-2} cm^{-1}) from the measured magnitudes using extra-atmosphere flux data taken from [12]. Observations made at nearby wavelengths were averaged. The logarithms of the observed flux values are shown by the circles in the figure.

In our spherically symmetrical model for the dust envelope of IRAS 18062 + 2410, we assumed that the envelope has a sharp inner boundary a distance r_1 from the center of the star, and is made up of identical spherical silicate grains with radius a = 0.10 micron. The number of grains per unit volume (concentration) $n(r)$

1063-7729/00/4410-0685$20.00 © 2000 MAIK "Nauka/Interperiodica"



Observed flux from the protoplanetary nebula IRAS 18062 + 2410 (circles) as a function of wavelength. The solid curve indicates the model spectral energy distribution for a star with a dust envelope.

varies in inverse proportion to the square of the distance from the star: $n(r) = n_1(r_1/r)^2$, where n_1 is the value at the inner boundary of the dust envelope. We took the radius of the outer boundary of the envelope to be $1000\,r_1$. For the complex dielectric function of the silicate grains, we used the recent data [13], based on IRAS observations of circumstellar dust envelopes. The complex absorption coefficient determined using these data was used to calculate the absorption-efficiency and scattering coefficients of the dust grains. The calculations were based on Mee theory and the code of Wiscombe [14]. The effective temperature $T_{\rm eff}$ and luminosity L of the central star determined in [5] are $T_{\rm eff}$ = 25000 K and $L = 10000\,L_\odot$. We assumed the spectral energy distribution to be a Planck distribution.

The equation of radiation transfer was solved using the CSDUST3 code. Our procedure is described in [15], and the code in [16]. We adopted a non-uniform wavelength grid of 59 points from 0.1 to 200 microns and a non-uniform radius grid of 45 points, with most readings concentrated near the inner envelope boundary. The distance to the star d, radius of the inner envelope boundary r_1, and envelope optical depth τ_{10} at λ = 10.0 microns were varied.

Multiple solutions of the problem yielded the following optimum parameter values: d = 6400 pc, $r_1 = 1.40 \times 10^{-3}$ pc, and τ_{10} = 0.050. The temperature of the dust grains at the inner envelope boundary was T_1 = 410 K. The spectral energy distribution of the model envelope is presented in the figure (solid curve). As we can see, it is in good qualitative agreement with the observed distribution. The data points show appreciable scatter, which does not, however, exceed that for dust-envelope models for other similar objects [1]. Discrepancies between the observed and model fluxes could have several origins. In our opinion, the most probable of these are the break of spherical symmetry of the envelope and differences between the true and adopted optical parameters of the dust. The derived distance slightly exceeds that obtained in [5] (d = 4800 pc); however, this is probably insignificant, given the uncertainties.

3. ESTIMATE OF POSSIBLE CONTRIBUTION OF NEBULA RADIATION

We can see from the figure that the observed flux systematically exceeds the model flux at far-IR wavelengths. This excess could be due to the contribution of the continuum radiation of the nebula surrounding the central star of IRAS 18062 + 2410. This nebula has already been detected in the spectrum of the object, from the presence of [OII], [NII], and [SII] forbidden lines; however, its excitation level is below the value required for the appearance of the nebular [OIII] lines characteristic of planetary nebulae. To estimate the flux of continuum radiation due to free–free and free–bound transitions, we use the parameters of the nebula obtained in [5]: electron density $N_e = 4 \times 10^4$ cm^{-3}, electron temperature T = 20000 K, and emission measure $N_e^2 V_n = 5.64 \times 10^{58}$ cm^{-3}, where V_n is the volume of the nebula. We adopt λ = 100 microns for the wavelength at which the contribution of free–free radiation is highest.

The power ε_ν per unit volume of a hydrogen nebula due to free–free transitions is [17]

$$\varepsilon_\nu = 5.4 \times 10^{-39} n_\nu g(\nu, T) N_e^2 T^{-1/2} \exp[-h\nu/kT], \quad (1)$$

where $\nu = c/\lambda$ is the frequency of the radiation, c the velocity of light, n_ν the index of refraction, $g(\nu, T)$ the Gaunt factor, h Planck's constant, and k Boltzmann's constant. Since the frequency corresponding to the chosen wavelength is much higher than the plasma frequency $\nu_p = 8.98 \times 10^5 N_e^{1/2}$ Hz, we can adopt n_ν = 1, and the Gaunt factor can be written [17]

$$g(\nu, T) = 0.54 \ln[5.0 \times 10^7 T^{3/2}/\nu]. \quad (2)$$

The observed flux $F(\lambda)$ is:

$$F(\lambda) = d^{-2} \int \varepsilon_\lambda dV, \quad (3)$$

where the integration is carried out over the entire volume of the nebula and $\varepsilon_\lambda = \varepsilon_\nu c/\lambda^2$.

Using the estimated electron temperature, we derive from (2) the Gaunt factor $g(\nu, T) = 2.1$. Using this value, substituting (1) into (3), and replacing the integral in (3) by the product of the square of the electron density and the nebula volume, we find that the maximum flux for a distance of d = 4800 pc [5] is 6.1×10^{-12} erg s^{-1} cm^{-2} cm^{-1}. This is only about 0.2% of the observed flux at 100 microns. Since the flux due to free–bound transitions must be much lower than that due to free–free transitions [17] at the given λ, we can completely neglect the effect of free–bound radiation.

Thus, the continuum radiation of the nebula cannot be responsible for the observed flux excess at far-IR wavelengths. Along with the origins noted above, this excess may be associated with the presence of another cool dust envelope ejected by the star in a previous stage of its evolution.

4. ESTIMATION OF STELLAR WIND PARAMETERS

When IRAS 18062 + 2410 was on the asymptotic giant branch, it should have had an intense stellar wind resulting in the loss of a large fraction of the mass of the star. For stars with low effective temperatures, the stellar wind is associated with radiation pressure on dust grains that have condensed in the outer layers of the extended stellar atmosphere, with a subsequent transfer of momentum to the gaseous medium. In this case, the velocity of the dust grains exceeds the velocity of expansion of the gas (the so-called drift velocity). The winds of hot stars are associated with radiation pressure on the gas, and the velocity of the dust grains (if they are able to condense) is lower than the velocity of motion of the gas. As the effective temperature of the core of a protoplanetary nebula increases, these mechanisms should follow one another. However, if the outer layers of the star are ejected rapidly enough, for a short time, the inner boundary of the dust envelope will reach a distance where the acceleration of the dust grains becomes essentially zero due to dilution of the radiation and decrease of the gas density. Thus, to a first approximation, we can consider the dust envelope of IRAS 18062 + 2410 to be quasi-steady-state until the time of ejection of the outer layers and to move inertially after this time.

It follows from the continuity equation that, in the case of a stationary dust envelope, the dust mass-loss rate $\dot{M}_d$ can be written

$$\dot{M}_d = 4\pi r_1^2 n_1 m_d V_d, \quad (4)$$

where $m_d = 4\pi a^3 \rho/3$ is the mass of a dust grain, ρ is the silicate density, and V_d is the velocity of motion of the dust. When Eq. (4) is used, some average value for V_d is generally adopted for a given group of objects. For example, it was assumed in the study [1] of models for protoplanetary nebulae that V_d = 10 km/s. For stellar wind driven by radiation pressure on dust with subsequent momentum transfer to the gas, a self-consistent procedure for determinating various parameters was suggested in [18]. However, this requires information about the gas expansion velocity at the inner envelope boundary and its velocity an infinite distance from the star. In the case of IRAS 18062 + 2410, these velocities are unknown. However, in the framework of our assumptions about the dust-envelope motion, we can estimate V_d using evolutionary calculations [6].

As was shown in [5], IRAS 18062 + 2410 lies near the evolutionary track for a star with initial mass 4 $M_\odot$ and final mass of the central star of the nebula 0.696 $M_\odot$. At the time when the star leaves the asymptotic giant branch, which can be assumed to coincide with the ejection of the envelope, its effective temperature is 7000 K and its luminosity is $L = 10000\ L_\odot$ [6]. The silicate condensation temperature is generally taken to be around 1000 K [7, 8].

Using the CSDUST3 code, we find that, for the given parameters for the dust and central star, a dust grain reaches the condensation temperature at a distance $r_0 = 0.15 \times 10^{-3}$ pc. As shown above, the present inner boundary of the dust envelope is at a distance $r_1 = 1.40 \times 10^{-3}$ pc. According to the evolutionary calculations of [6], after its exit from the asymptotic giant branch, the central star heats to its present temperature, $T_{\rm eff}$ = 25000 K over $t \approx 100$ yrs. Thus, we can adopt as an estimate of the expansion velocity for the dust envelope the value $V_d = (r_1 - r_0)/t$, which, in our case, is 12 km/s. Since r_0 is almost an order of the magnitude smaller than r_1, even large variations of the stellar parameters and condensation temperature should not appreciably affect V_d.

The resulting value for the velocity of motion of the dust is consistent with other independent V_d estimates for silicate dust envelopes [1, 7–9]. Substituting this value in Eq. (4), taking ρ = 3.5 g/cm^3 for the silicate density, and using the derived dust-grain concentration $n_1 = 2.7 \times 10^{-7}$ cm^{-3}, we find that the dust mass-loss rate is $\dot{M}_d = 1.8 \times 10^{-8} M_\odot$/yr. According to the data of [8], the ratio of the loss of gas $\dot{M}_g$ and dust $\dot{M}_d$ mass for the silicate dust envelopes of stars is approximately 250. Thus, the total mass-loss rate for IRAS 18062 + 2410 in the stage of evolution directly preceding the envelope's ejection is inferred to be $\dot{M} = 4.5 \times 10^{-6} M_\odot$/yr.

5. CONCLUSIONS

We have shown that, at the time of ejection of the envelope, the protoplanetary nebula IRAS 18062 + 2410 was among those objects having the highest known mass-loss rates. However, stellar evolutionary calculations indicate that the mass-loss rate could have been even higher before the envelope ejection, reaching about 10^{-4} $M_\odot$/yr at the tip of the asymptotic giant branch [6]. As a result, the density distribution of the dust envelope surrounding the object should depend on the previous mass-loss history. It is difficult to predict this distribution in advance. It is possible that the observed far-IR excess is primarily due to a deviation of the true density distribution from that adopted in the model. The deviations of the measured fluxes from the model values could also be due to differences between the adopted and true optical parameters of the dust and possible violation of spherical symmetry of the envelope.

Nevertheless, our calculated model describes the observed flux quite adequately over a very wide range of wavelengths. We have obtained an independent estimate of the dust velocity within the framework of our assumptions, which is in good agreement with values from other studies. Thus, we believe that the derived parameters of the dust envelope are close to their real values.

ACKNOWLEDGMENTS

This work was supported by the State Science and Technology Program "Astronomy" and the Ministry of Education of the Russian Federation.

REFERENCES

1. K. M. Volk and S. Kwok, Astrophys. J. **342**, 345 (1989).
2. G. Lawrence, T. J. Jones, and R. D. Gehrz, Astron. J. **99**, 1232 (1990).
3. R. D. Oudmaijer, W. E. C. J. van der Veen, L. B. F. M. Waters, *et al.*, Astron. Astrophys., Suppl. Ser. **96**, 625 (1992).
4. P. García-Lario, A. Manchado, W. Pych, and S. R. Pottasch, Astron. Astrophys., Suppl. Ser. **126**, 479 (1997).
5. V. P. Arkhipova, N. P. Ikonnikova, R. I. Noskova, *et al.*, Pis'ma Astron. Zh. **25**, 30 (1999) [Astron. Lett. **25**, 25 (1999)].
6. T. Blocker, Astron. Astrophys. **299**, 755 (1995).
7. P. J. Bedijn, Astron. Astrophys. **186**, 136 (1987).
8. W. A. Schutte and A. G. G. M. Tielens, Astrophys. J. **343**, 369 (1989).
9. O. Hashimoto, Y. Nakada, T. Onaka, *et al.*, Astron. Astrophys. **227**, 465 (1990).
10. D. Burstein and C. Heiles, Astron. J. **87**, 1165 (1982).
11. C. W. Allen, *Astrophysical Quantities* (Athlone, London, 1973; Mir, Moscow, 1977).
12. V. L. Straĭzhis, *Multiwavelength Photometry of Stars* (Mokslas, Vilnius, 1977).
13. P. David and B. Pegourie, Astron. Astrophys. **293**, 833 (1995).
14. W. Wiscombe, Appl. Opt. **19**, 1505 (1980).
15. C. M. Leung, J. Quant. Spectrosc. Radiat. Transfer **16**, 559 (1976).
16. M. P. Egan, C. M. Leung, and G. F. Spagna, Comput. Phys. Commun. **48**, 271 (1988).
17. K. R. Lang, *Astrophysical Formulae: a Compendium for the Physicist and Astrophysicist* (Springer, Heidelberg, 1974; Mir, Moscow, 1978), Part 1.
18. M. A. T. Groenewegen, P. A. Whitelock, C. H. Smith, and F. Kerschbaum, Mon. Not. R. Astron. Soc. **293**, 18 (1998).

Translated by K. Maslennikov

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 689–695. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 777–783.

Rapid *UBVRI* Photometry of the Active Flare Stars EV Lac and AD Leo

I. Yu. Alekseev, V. E. Chalenko, D. N. Shakhovskoĭ

Crimean Astrophysical Observatory, p/o Nauchnyĭ, Crimea, 334413 Ukraine

Received October 15, 1999

Abstract—Regular photometric observations of the two active flare stars EV Lac and AD Leo were carried out in 1996–1999. We have compared the location of the intrinsic radiation of the strongest flares in *UBVRI* two-color diagrams with those predicted by various models for the radiation sources. Our statistical analysis of flares in EV Lac confirms the 7.5-year activity cycle reported earlier. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

EV Lac (dM4.5e) and AD Leo (dM3.5e) are among the best-studied active flaring red dwarf stars, and display a wide range of photometric variability on a variety of time scales, from short-lived flares (beginning with 0.1 s) to slow BY Dra-type variability. Regular, coordinated observations of EV Lac at various wavelengths have been conducted at the Crimean Astrophysical Observatory since 1986 [1–7]. Earlier, we monitored AD Leo during a search for activity at decameter wavelengths [8]. In all these studies, we achieved a time resolution of 19 s. Here, we continue our photometric investigation of the flare activity of EV Lac and AD Leo with significantly better time resolution (0.1 s).

2. OBSERVATIONS

All observations were carried out using the AZT-11 1.25-m reflector with the double-image chopping *UBVRI* photometer/polarimeter designed by Piirola [9] and adapted for time-resolution photometry [10, 11]. The time resolution achieved was 0.1 s. We used a diaphragm with a 20″ diameter. The sky background was not measured simultaneously (as in the usual regime), but instead separately, during the two to three minutes ending each hour of monitoring. A comparison star was observed after the sky background measurements. The comparison stars for EV Lac and AD Leo were SAO 52337 ($V = 9\overset{m}{.}19$, $U–B = 1\overset{m}{.}02$, $B–V = 1\overset{m}{.}18$, $V–R = 0\overset{m}{.}91$, $V–I = 1\overset{m}{.}39$ [12]) and for AD Leo BD + 20°2468 ($V = 10\overset{m}{.}24$, $U–B = 0\overset{m}{.}06$, $B–V = 0\overset{m}{.}44$, $V–R = 0\overset{m}{.}43$, $V–I = 0\overset{m}{.}72$ [13]), respectively.

We observed AD Leo on five nights from May 10 to May 21, 1997 and on five nights in April 1999. The 1999 observations were made within the framework of an international coordinated program [14]. The total durations T of the 1997 and 1999 observations were 9.93 and 11.5 hours, respectively. Four flares were recorded. Table 1 presents their general characteristics: the date and time of flare maximum; duration of flare onset T_b and decay T_a, in minutes; equivalent flare duration $P = \int[(I_{fl} – I_0)/I_0]dt$; and flare amplitude ΔU. The total equivalent duration of the flares was $\Sigma P = 45.91$ min, with the corresponding formal activity index of the star $\Sigma P/T$ equal to 0.052 in 1997 and to 0.021 in 1999, which is quite typical for AD Leo [15].

We observed EV Lac during the four nights of October 7 and 15–17, 1996 (Table 2) and on December 3, 1996. The total observation time was 6.83 h. Five flares were recorded, including a very strong flare on October 15, 1996 at UT 20:44. The activity index was 0.350, the highest ever observed for EV Lac. Because of the short total observing time, this value is not statistically significant.

Table 1. Flares of AD Leo

Date	UT	T_b, min	T_a, min	P, min	ΔU
May 18, 1997	18:48:09	0.9	21.2	17.18	$2\overset{m}{.}26$
	19:10:16	0.7	19.3	11.42	2.46
May 19, 1997	18:55:53	0.3	8.6	2.51	1.49
Apr. 30, 1999	19:37:41	2.8	5.0	14.80	0.60

Table 2. 1996 flares of EV Lac

Date	UT	T_b, min	T_a, min	P, min	ΔU
Oct. 7, 1996	23:50:17	0.1	1.2	0.16	$0\overset{m}{.}50$
Oct. 15, 1996	19:50:42	1.6	7.9	0.57	0.20
	20:44:27	3.7	>20	>142	3.73
Oct. 16, 1996	21:23:11	3.1	9.0	0.60	0.15
	22:05:05	0.3	1.5	0.10	0.18

1063-7729/00/4410-0689$20.00 © 2000 MAIK "Nauka/Interperiodica"

Table 3. 1997 flares of EV Lac

Date	UT	T_b, min	T_a, min	P, min	ΔU
Aug. 22, 1997	20:36:35	0.4	6.2	1.00	$0^m.34$
Aug. 23, 1997	20:19:48				0.44
	25:25:47	0.5	3.7	0.75	0.56
Aug. 26, 1997	24:37:30	0.7	2.9	0.32	0.24
	24:46:42	0.3	2.2	0.57	0.38
Aug. 28, 1997	22:34:19	1.5	2.2	1.55	0.64
Aug. 29, 1997	22:39:17	0.3	3.2	0.89	0.76
	23:48:11	6.6	29.3	20.37	1.84
Oct. 6, 1997	18:03:54	0.1	0.4	0.14	0.49
	18:43:36	0.2	0.3	0.22	0.52
Oct. 7, 1997	17:40:37	0.1	0.3	0.10	0.29
	19:39:00	1.3	4.0	0.70	0.25

Table 4. 1998 flares of EV Lac

Date	UT	T_b, min	T_a, min	P, min	ΔU
Sep. 10, 1998	20:30:37	0.4	6.5	4.06	$1^m.58$
	23:48:53	0.2	2.2	0.41	0.56
	24:03:05	0.1	1.6	0.17	0.45
	24:12:06	0.5	3.3	1.19	0.95
Sep. 11, 1998	18:58:16	4.7	14.1	6.55	1.03
	19:37:24	0.5	3.1	0.68	0.72
	19:53:20	5.7	20.9	5.00	0.34
	21:55:09	3.1	86.2	41.13	2.00
	25:05:00	0.2	1.1	0.14	0.20
	25:19:36	0.5	3.3	0.55	0.68
Sep. 12, 1998	20:02:05	2.0	2.9	1.00	0.31
Sep. 14, 1998	19:08:08	0.6	2.4	0.89	0.57
Oct. 10, 1998	19:14:36	3.1	40	54.69	2.72

In 1997, we observed EV Lac during seven nights from August 22–31 (21.60 h, eight flares) and on three nights from October 6–9 (8.81 h, four flares, $\Sigma P/T$ = 0.015). The main parameters of the flares are given in Table 3.

In 1998, we observed EV Lac on September 10–12 and 14–16, and also on October 10. The total observing time was 36.35 h, during which 13 flares were detected, with a total equivalent duration of 116.46 min. The data for these flares are collected in Table 4.

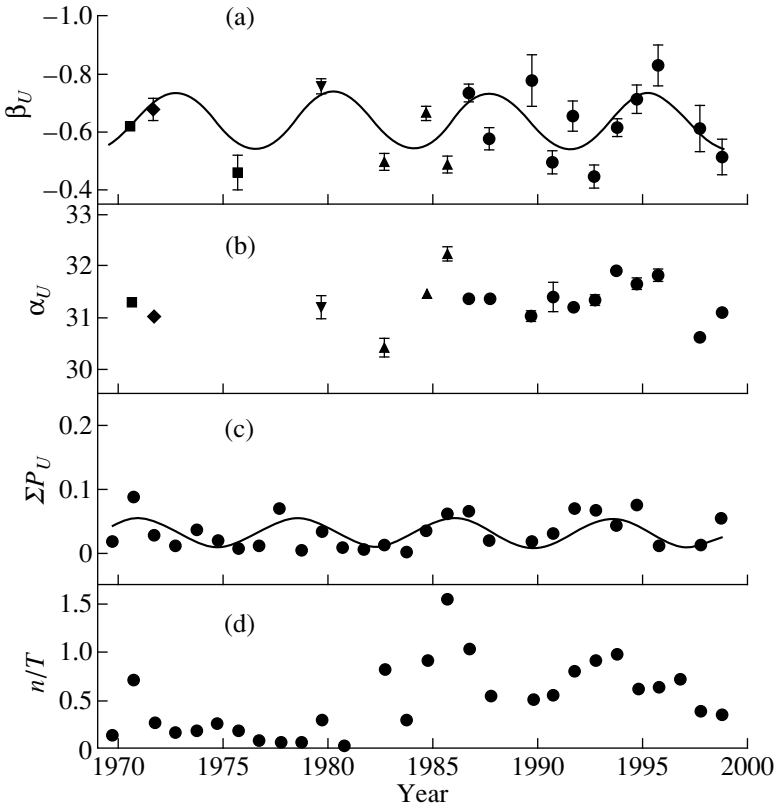


Fig. 1. Statistical parameters of the flares: (a) spectral index, β_U; (b) logarithm of the expected maximum energy, α_U; (c) index of flare activity, $\Sigma P_U/T$; and (d) mean flare rate, n/T.

Table 5. Intrinsic radiation of strong flares of EV Lac

Date	UT	U_{max}	*U–B*	*B–V*	*V–R*	*V–I*
Oct. 15, 1996	20:44:27	$9^m.15 \pm 0^m.02$	$-0^m.78 \pm 0^m.06$	$0^m.02 \pm 0^m.10$	$0^m.51 \pm 0^m.14$	$0^m.55 \pm 0^m.22$
May 18, 1997	18:48:09	9.27 ± 0.06	−1.19 ± 0.15	0.09 ± 0.28		
May 18, 1997	19:10:16	9.39 ± 0.06	−0.87 ± 0.14	−0.30 ± 0.28		
May 19, 1997	18:55:53	10.60 ± 0.10	−0.88 ± 0.26			
Sep. 11, 1998	21:55:09	10.75 ± 0.07	−0.88 ± 0.20	0.21 ± 0.38		
Oct. 10, 1998	19:14:36	10.28 ± 0.02	−1.03 ± 0.05	0.08 ± 0.11	0.29 ± 0.18	0.12 ± 0.33
Apr. 30, 1999	19:37:41	11.60 ± 0.02	−0.92 ± 0.03	0.14 ± 0.06	0.63 ± 0.14	

3. STATISTICS OF THE FLARES

The most commonly used statistical parameters of stellar flare activity are the mean flare rate n/T and the activity index $\Sigma P_U/T$. Their analysis of these two parameters for AD Leo led Pettersen and Panov [15] to suspect an eight-year cycle in the star's activity. Our data (four flares in 23 h of observations) are obviously insufficient to test this result.

The flare rate for EV Lac in 1996 was 0.73 h^{-1}, in full consistency with the Crimean data for previous years (0.51–1.04 h^{-1} [7]). The flare rates in 1997 and 1998 were 0.39 and 0.36 h^{-1}, respectively, the lowest during all 12 years of Crimean observations of this star.

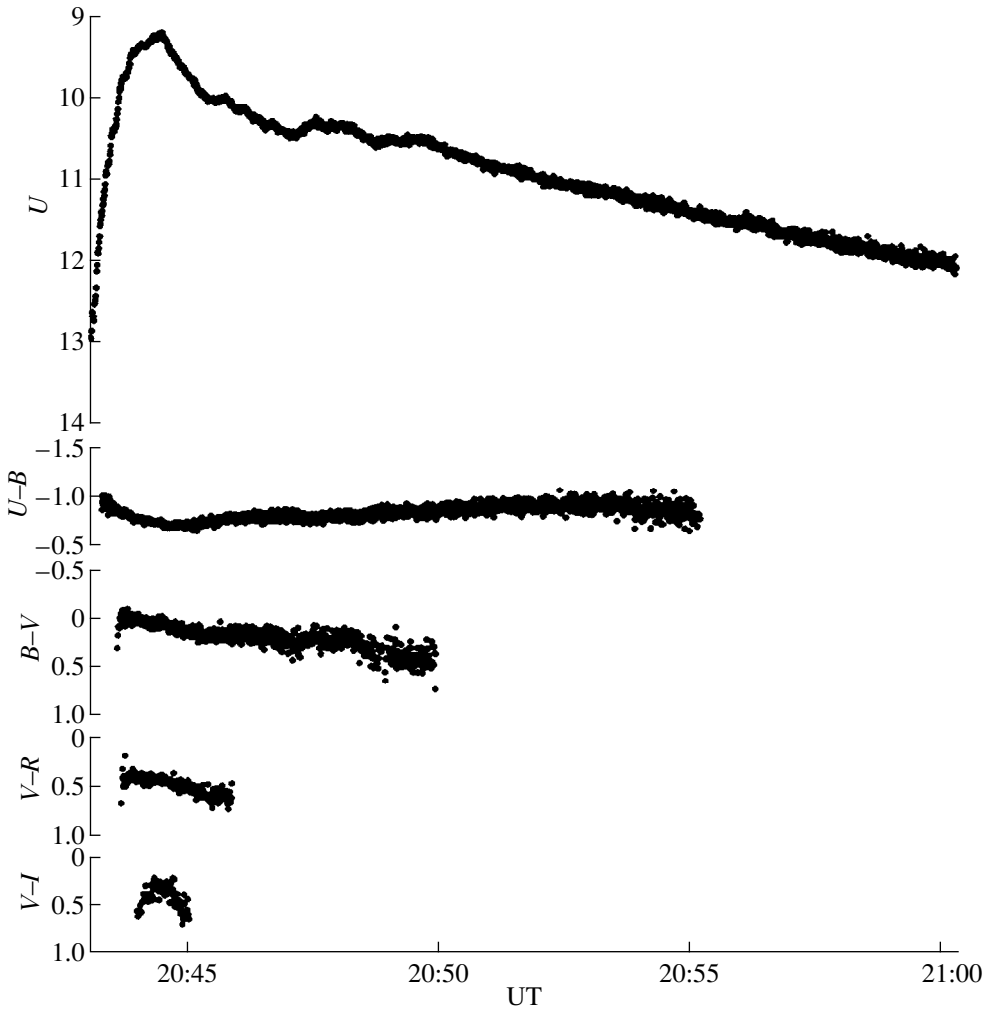


Fig. 2. Light and color curves of the October 15, 1996 UT 20:44 flare of EV Lac.

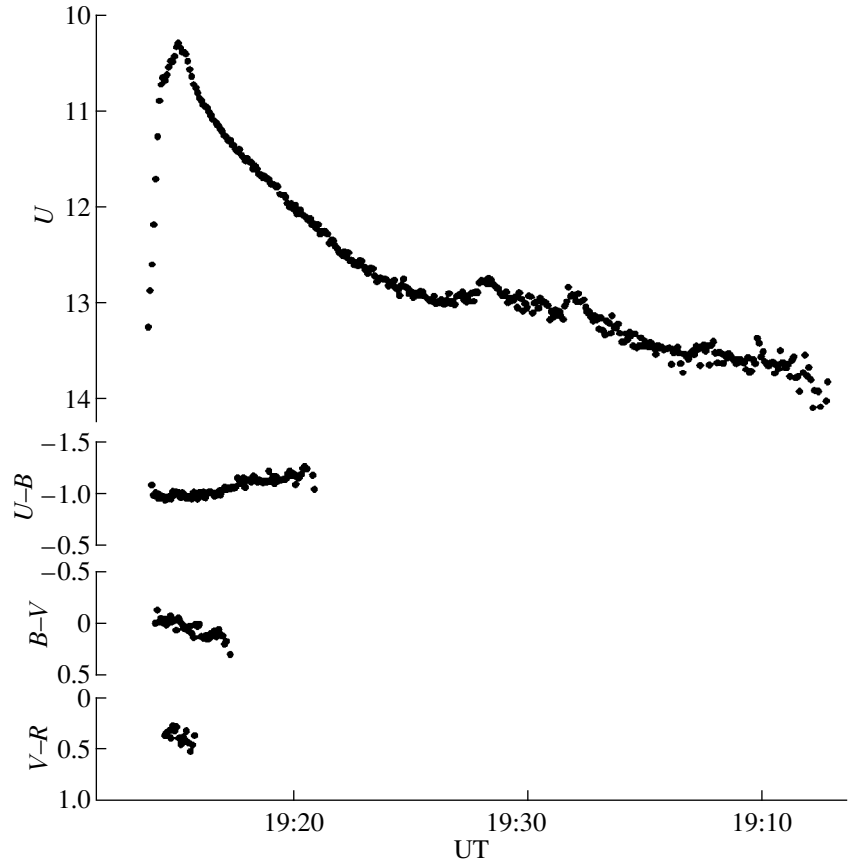


Fig. 3. Light and color curves of the October 10, 1998 UT 19:14 flare of EV Lac.

During the entire period of photoelectric observations of EV Lac since 1969, the lowest flare rate was 0.05 h^{-1}. This value shows no cyclic behavior (Fig. 1d), but note that n/T is very sensitive to the instrumentation used, and is statistically trustworthy only when homogeneous observations are considered.

The product of the activity index $\Sigma P_U/T$ and the star's *U*-band luminosity (for EV Lac, this is $E_U = 3.75 \times 10^{30}$ erg/min [7]) describes the mean energy release in flares during the observation interval. This index was obviously overestimated for 1996 due to the occurrence of a single strong flare during a short total observing time; it was 0.014 and 0.053 in 1997 and 1998, respectively. Alekseev and Gershberg [7, 16] showed that this quantity showed cyclic behavior, with a typical cycle duration of seven to eight years. Our 1997 and 1998 observations are in good agreement with their relation (Fig. 1c).

The most complete statistical analysis of the flares' optical radiation uses the so-called energy spectrum—the dependence of the rate of the flares on their energy (see, for example, [17–19]). Here, we plot the logarithm of the total *U*-band radiation of the flares as a function of the logarithm of the so-called accumulated flare rate, $\nu(E_U) = (\Sigma n)/T$, where the sum is over all flares with energies E_U and above. At moderate flare energies, this relation corresponds to a straight line of the form $\log\nu(E_U) = \text{const} + \beta_U \log E_U$. The slope of this line, β_U, was 0.62 ± 0.08 in 1997 and 0.52 ± 0.06 in 1998. These spectral indices are also cyclic, with a 7–8 year time scale, and the values found by us are in good agreement with the relation of Alekseev and Gershberg [7, 16].

Another parameter of the flare energy spectrum is the logarithm of the expected maximum flare energy during a 10-hour α_U monitoring session. The values found by us were 30.63 ± 0.04 in 1997 and 31.10 ± 0.07 in 1998, in agreement with the data for other years (Figs. 1a, 1b).

4. ANALYSIS OF STRONG FLARES

We considered seven strong flares for which we were able to derive the colors of the pure flare radiation in more detail. Papers [1–6] describe in detail how to derive and analyze these colors. The brightnesses and colors of the intrinsic radiation at flare maximum are collected in Table 5. Our comparison with earlier observations of EV Lac shows that the color indices at the

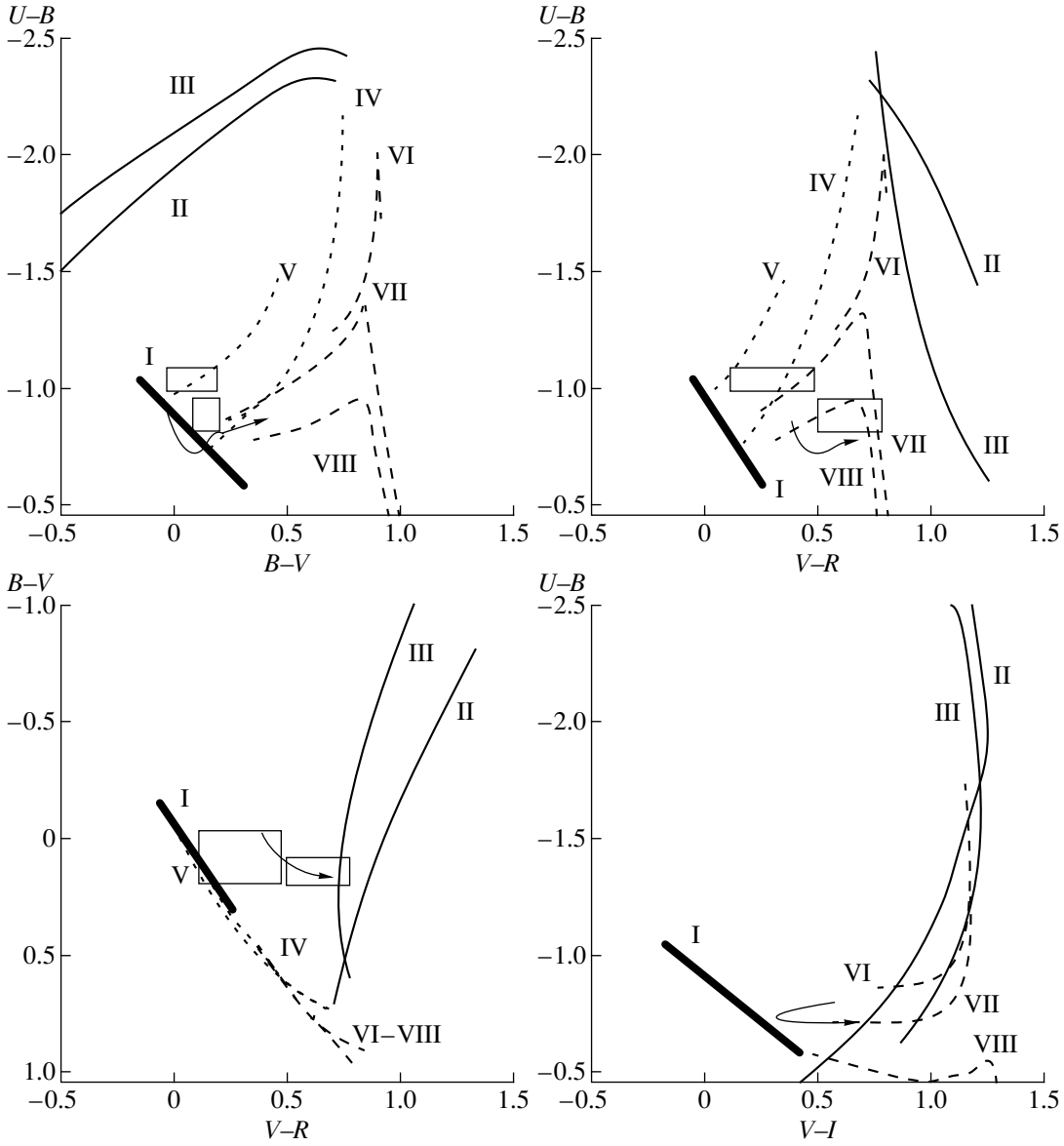


Fig. 4. Two-color diagram for changes of the intrinsic radiation of the October 15, 1996, strong flare of EV Lac (heavy solid) and for the maximum radiation of the October 10, 1998, flare of EV Lac and April 30, 1999, flare of AD Leo (rectangles). The dashed curves are discussed in the text.

maximum phases of the strong flares do not differ from those for flares observed earlier.

We were able to follow the time behavior of the pure flare radiation in all four colors fairly reliably only for the October 15, 1996 UT 20:44 flare of EV Lac. For the less strong flare of EV Lac on October 10, 1998 at UT 19:14, we traced the behavior of three color indices: *U–B*, *B–V*, and *V–R*. Figures 2 and 3 show the light and color curves of the flares. We rejected any brightnesses and color indices with quantum statistical errors in excess of $0\overset{m}{.}25$, and they are not displayed in the figures. None of the flares showed any short-term brightness or color variations. The *U–B* color index reddened on the ascending branch of the October 15, 1996 flare, from $-1\overset{m}{.}0$ to $-0\overset{m}{.}6$, and then returned to its original value. *B–V* and *V–R* both reddened, from approximately $0\overset{m}{.}0$ to $0\overset{m}{.}3$ and from $0\overset{m}{.}4$ to $0\overset{m}{.}7$, respectively. *V–I* became bluer rather rapidly on the ascending branch, and then, also rapidly, began to become redder.

The *U–B* values for the October 10, 1998 flare are generally bluer than for the first flare, whereas *B–V* and *V–R* are similar. As for the first flare, *U–B* reddened on the ascending branch, to $-1\overset{m}{.}0$, and then began to slowly become bluer, whereas *B–V* and *V–R* reddened.

Figure 4 shows (*U–B*, *B–V*), (*U–B*, *V–R*), (*U–B*, *V–I*), and (*B–V*, *V–R*) two-color diagrams displaying the color behavior of the October 15, 1996, flare of EV Lac (thick solid lines) and the color indices at maximum brightness for the strong flares of EV Lac on October 10, 1998, and of AD Leo on April 30, 1999 (rectangles). The color behavior of the October 15, 1996, flare was approximated with a polynomial fit.

In addition to the flare colors, the diagrams show curves for blackbodies with temperatures from 8000 to 20000 K (I); hydrogen plasma transparent in the Balmer continuum for a temperature of 10000 K and electron densities of 10^{12} (II) and 10^{14} (III) cm^{-3}; plasma optically thick in the Balmer continuum for T_e = 10000 K (IV) and T_e = 15000 K (V); and radiation from the upper layers of the star's atmosphere heated by a stationary flow of high-energy protons at threshold energies of 1 (VI), 3 (VII), and 5 (VIII) MeV. Descriptions of the computation techniques used to generate the two-color diagrams for the various models can be found in [20–23].

In the *U–B*, *B–V* diagram, the premaximum radiation of the October 15, 1996, flare is located near the blackbody curves and shifts toward cooler regions (from 20000 to 10000 K) as the flare maximum is approached. As the flare decays, the colors move toward the right, toward model VII. In the *U–B*, *V–R* diagram, the color curve is displaced to the right of the region of blackbody radiation along the *V–R* axis, into the region of proton heating models. The colors in the *B–V*, *V–R* diagram are also shifted to the right and are located between the blackbody curves and the region of optically thin plasma radiation (III). In the course of time, we move toward the bottom right in the diagram. In the *U–B*, *V–I* diagram, the flare radiation is located in the region of the proton heating model (VII), and approaches the region of blackbody radiation at maximum brightness. The positions of the flare maxima for the weaker flares of October 10, 1998, and April 30, 1999, coincide with the colors of the strong flare within the errors. Thus, the near-maximum flare radiation is close to blackbody in all colors but *R*. This location for the flare radiation in the two-color diagrams is also typical of other flares on the same stars [1–6]; as in previous studies, it does not correspond purely to any of the discussed sources of intrinsic flare radiation.

6. CONCLUSIONS

In our analysis of multicolor photoelectric observations of the active stars EV Lac and AD Leo, we have not considered manifestations of their activity on very short time scales (below 1 s). Detection and analysis of such rapid, low-amplitude activity requires special mathematical techniques [24–26], and we will investigate this separately. All the flares detected and considered by us here had development time scales from several tens of seconds to minutes, and were subject to a standard analysis. Our statistical analysis of the flares of EV Lac has confirmed the existence of cyclic behavior with a 7- to 8-year period, initially reported by Alekseev and Gershberg [7, 16].

We have considered the intrinsic flare radiation and detected nonmonotonic variations in the flare color indices, in agreement with the results of earlier observations. Near maximum brightness, the flare radiation is close to that of blackbodies with temperatures from 10000 to 20000 K in all colors but *R*. The intrinsic flare radiation does not correspond purely to any of the discussed possible sources of radiation.

REFERENCES

1. R. E. Gershberg, V. P. Grinin, I. V. Il'in, *et al.*, Astron. Zh. **68**, 548 (1991) [Sov. Astron. **35**, 269 (1991)].
2. R. E. Gershberg, I. V. Il'in, A. N. Rostopchina, *et al.*, Astron. Zh. **70**, 984 (1993) [Astron. Rep. **37**, 497 (1993)].
3. I. Yu. Alekseev, R. E. Gershberg, I. V. Il'in, *et al.*, Astron. Astrophys. **288**, 502 (1994).
4. H. Abdul-Aziz, E. P. Abranin, I. Yu. Alekseev, *et al.*, Astron. Astrophys., Suppl. Ser. **114**, 509 (1995).
5. E. P. Abranin, L. L. Bazelyan, I. Yu. Alekseev, *et al.*, Astrophys. Space Sci. **257**, 131 (1997–1998).
6. E. P. Abranin, I. Yu. Alekseev, S. Avgoloupis, *et al.*, Astron. Astrophys. Trans. **17**, 221 (1998).
7. I. Yu. Alekseev and R. E. Gershberg, in *The Earth and the Universe*, Ed. by G. Asteriadis, A. Bantelas, M. Contadakis, *et al.* (Univ. of Thessaloniki, 1997), p. 43.
8. E. P. Abranin, I. Yu. Alekseev, L. L. Bazelyan, *et al.*, Kinematika Fiz. Nebesnykh Tel. **10**, 70 (1994).
9. V. Piirola, Report No. 6 (Observatory of Astrophysical Laboratory, Univ. of Helsinki, 1984), p. 151.
10. S. Yu. Kalmin, Kinematika Fiz. Nebesnykh Tel **11** (3), 82 (1995).
11. S. Yu. Kalmin and D. N. Shakhovskoĭ, Kinematika Fiz. Nebesnykh Tel **11** (3), 85 (1995).
12. I. Yu. Alekseev, Astron. Zh. **73**, 81 (1996) [Astron. Rep. **40**, 69 (1996)].
13. I. Yu. Alekseev, Astron. Zh. **75**, 736 (1998) [Astron. Rep. **42**, 649 (1998)].
14. M. Güdel, M. Audard, E. F. Guinan, *et al.*, in *Cool Stars, Stellar Systems and the Sun*, Ed. by R. J. García López, R. Rebolo, and M. R. Zapatero Osorio, Astron. Soc. Pac. Conf. Ser. (in press).
15. B. R. Pettersen and K. P. Panov, in *Cool Stars, Stellar Systems and the Sun*, Ed. by M. Zeilik and D. M. Gibson, Lect. Notes Phys. (Springer, Berlin, 1986), Vol. 254, p. 91.
16. I. Yu. Alekseev and R. E. Gershberg, in *Cool Stars, Stellar Systems and the Sun*, Ed. by R. A. Donahue and J. A. Bookbinder, Astron. Soc. Pac. Conf. Ser. **154**, 1471 (1999).

17. R. E. Gershberg, Astrofizika **22**, 531 (1985).
18. R. E. Gershberg, Mem. Soc. Astron. Ital. **60**, 263 (1988).
19. R. E. Gershberg, in *Eruptive Stars*, Ed. by A. G. Masevich (MID, Moscow, 1996), p. 189.
20. V. P. Grinin and N. A. Katysheva, Izv. Krym. Astrofiz. Obs. **62**, 66 (1980).
21. V. P. Grinin and V. V. Sobolev, Astrofizika **31**, 527 (1989).
22. V. P. Grinin, V. M. Loskutov, and V. V. Sobolev, Astron. Zh. **70**, 350 (1993) [Astron. Rep. **37**, 182 (1993)].
23. N. N. Chalenko, Astron. Zh. **76**, 529 (1999) [Astron. Rep. **43**, 459 (1999)].
24. B. E. Zhilyaev, Kinematika Fiz. Nebesnykh Tel **10** (1), 32 (1994).
25. B. E. Zhilyaev, Ya. O. Romanyuk, and O. A. Svyatogorov, Kinematika Fiz. Nebesnykh Tel **10** (6), 88 (1994).
26. B. E. Zhilyaev, I. A. Verlyuk, Ya. O. Romanyuk, *et al.*, Astron. Astrophys. **334**, 931 (1998).

Translated by N. Samus'

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 696–700. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 784–789.

Statistics of BY Draconis Variables

I. Yu. Alekseev

Crimean Astrophysical Observatory, p/o Nauchnyĭ, Crimea, 334413 Ukraine

Received June 22, 1999

Abstract—A statistical survey of 113 spotted red dwarf stars that are known or suspected BY Draconis variables is presented. Typical indicators of stellar photometric activity—the amplitudes of the rotational modulation and seasonal mean brightness variations ΔV and $\Delta\langle V\rangle$—are compared to the global parameters of the stars. First, photometric variability shows a weak dependence on spectral type; second, ΔV and $\Delta\langle V\rangle$ grow with increasing stellar rotational velocity and decreasing Rossby number, with the dependences saturating at the critical values $V_{crit} \sim 15$–20 km/s and $Ro_{crit} \sim 0.2$–0.3; and third, the Sun as a star fits well into the derived relations. Thus, the spottedness of stars, like other indicators of stellar activity, depends on their global parameters.

1. INTRODUCTION

Starspots are a common phenomenon for cool stars whose convective envelopes play a substantial role in determining their internal structures. These are primarily emission-line red dwarfs, whose activity is qualitatively similar to that of the Sun: BY Dra variables. T Tauri and post-T Tauri stars; rapidly rotating, presumably single, FK Com-type G giants; RS CVn binary systems consisting of an F–G subgiant or dwarf and a G–K subgiant; W UMa close binary systems; and cool secondary components of some Algol and symbiotic systems with spots are also known.

Stars of all these types are photometric variables: the axial rotation of a spotted star leads to a low-amplitude, periodic rotational brightness modulation. In addition, changes in the configuration of the spotted regions lead to variations in the light curve's shape and fluctuations in the star's mean brightness, the amplitude of the rotational modulation, and the phase of minimum brightness.

A wide range of manifestations of solar activity are observed for BY Dra stars: evidence for powerful chromospheres and coronas, magnetic fields, UV Ceti-type flares, and other phenomena. The dependences of these manifestations on spectral type, rotational velocity, and Rossby number are well known (see, for instance, [1–6]). On the other hand, the Sun as a star can also be considered a low-activity BY Dra variable. Consequently, a new branch of astronomy, solar–stellar physics, has emerged. As part of this field, studies of spotted red dwarfs can improve our knowledge of stellar magnetism, rotation, stellar gas-dynamics processes, and other physical properties of stars. Such studies can also be used to generalize solar activity patterns to stars of different masses and ages.

High-resolution spectroscopic observations are very informative for studies of starspot activity, making it possible to follow variations of line profiles due to spots over the course of the star's period of axial rotation. However, with rare exceptions, red dwarfs are usually too faint for such studies. For this reason, they have primarily been studied photometrically. In the current study, we present a list of all published photometric observations of BY Dra stars (more than 300 papers) and study the dependences of their variability on basic stellar parameters.

The first compilation of photometric observations presented by Alekseev and Gershberg [7] included 70 BY Dra stars. In the present study, we expand this list to 113 objects. Most of these stars are UV Ceti flare stars from various catalogs (for example, [8, 9]). Most of the binary spotted stars in our sample are also in the catalogs of chromospherically active binaries of Strassmeier *et al.* [10, 11]. We have tentatively grouped the stars into two lists: reliable (well-studied) and poorly studied variables (candidates). For each of the lists, we compiled three tables (Tables 1–3 and 4–6, respectively); these are available electronically.

Electronic Tables 1 and 4, "Identifications," contain the GCVS names of the stars; their numbers according to the Gliese, HD, BD/CoD, SAO, and other catalogs; and their coordinates for the equinox 1950.0. Electronic Tables 2 and 5, "Main Properties," contain the stars' spectral types; absolute magnitudes M_V; U–B, B–V, V–R, and V–I color indices; projected rotational velocities $V\sin i$; and information about their binary natures. We tried to give photometric values for the maximum brightness state, when the starspot level is minimum. We expect that, for well-studied stars, these values characterize the unspotted photospheres.

Electronic Tables 3 and 6, "Photometric Variability," contain the stars' maximum brightness levels outside flares, largest amplitudes of rotational brightness modulation ΔV, amplitudes of the mean brightness variations due to changing spot configuration $\Delta\langle V\rangle$, number of observed epochs, photometric periods, references to

1063-7729/00/4410-0696$20.00

papers describing the variability, and remarks. The remarks indicate whether a variable is an active flare star (F), solar-type star (S), possible RV CVn variable (R), or possible spotted T Tauri or post-T Tauri star (T). The list of candidates also includes a number of poorly studied stars whose spot activity is suspected based on automated survey observations [12, 13] (A) or observations of stars whose activity was detected by the EXOSAT X-ray telescope (E) [14–16].

2. THE STATISTICS OF BY DRA STARS

BY Dra stars are cool variables spanning the spectral range from G0Ve (HN Peg), where they merge with solar-type stars, to dM6e (HH And). Figure 1 shows the Hertzsprung–Russell diagram for all the stars in our sample, with different symbols for reliable, well-studied variables (solid circles), solar-type candidates and related objects (open circles), possible post-T Tauri stars (triangles), active flare stars with suspected spots (dots), and other red dwarfs whose spot activity was revealed by various survey observations (diamonds). This figure shows that all the BY Dra variables except for FK Ser, which is probably a T Tauri star, are on or slightly above the main sequence.

Thus, we will take BY Dra stars to be simply a subtype of flare star. This definition clearly separates BY Dra stars from other types of cool spotted stars, but is by no means the only one possible: for example, Hall [17] sets BY Dra stars apart from UV Ceti variables and solar-type stars, while Chugaĭnov [18] combines spotted red dwarfs into a single class with RS CVn systems and spotted T Tauri stars. As a subtype of flare star, spotted red dwarfs are the most widespread type of variable in our Galaxy, and include the Sun. According to our sample data, G stars make up 20% of the total number of known variables, whereas K and M stars each make up 40%. Among flare stars, however, M stars constitute more than half of all variables, and one-fourth of all variables are K stars. This difference is mainly due to observational selection effects. We cannot say anything definite about whether the decrease in spot activity toward late spectral types is real or not.

BY Dra stars show two types of photometric variability: quasi-sinusoidal, periodic rotational modulations and slow changes of the mean brightness due to changes in the configuration of the spotted regions. One hundred seven objects are known to display rotational brightness modulations. For one-third of the stars, the maximum amplitude ΔV is from $0\overset{m}{.}05$ to $0\overset{m}{.}10$, and for the vast majority of the stars (80%) it does not exceed $0\overset{m}{.}15$. However, ten stars showed rotational modulation amplitudes in excess of $0\overset{m}{.}20$, and for three variables—CC Eri (K7Ve), VY Ari (G9Ve), and AU Mic (dM2e)—$\Delta\langle V\rangle$ reached $0\overset{m}{.}30$. Slow changes of the mean brightness were seen in 89 stars. Usually (70% of the stars), they also do not exceed $0\overset{m}{.}15$, but the maximum $\Delta\langle V\rangle$ values for 16% of the variables were greater than $0\overset{m}{.}20$, and three stars—V833 Tau (K5Ve), V775 Her (K0Ve), and BY Dra (K6Ve)—have $\Delta\langle V\rangle$ amplitudes exceeding $0\overset{m}{.}30$, namely $0\overset{m}{.}31$, $0\overset{m}{.}43$ and $0\overset{m}{.}38$, respectively. When photographic observations are included, the variability of the mean brightness, Δm_{pg}, reaches $0\overset{m}{.}9$. 84 for V833 Tau and PZ Mon (K2Ve). Both signatures of BY Dra-type variability are present in 84 stars. For these variables, the dependence between the two brightness amplitudes is weak: the corresponding correlation coefficient is $r(\Delta V\Delta\langle V\rangle) = 0.60 \pm 0.13$.

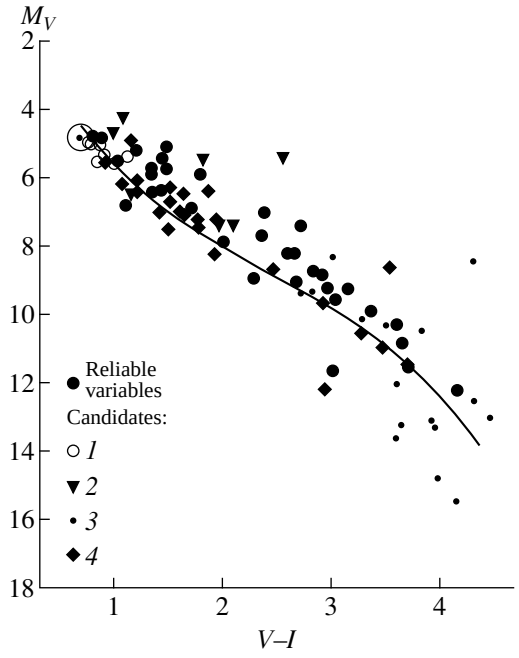


Fig. 1. Hertzsprung–Russell diagram for BY Dra stars and related objects: (*1*) solar-type stars, (*2*) spotted post-T Tauri stars, (*3*) flare stars, and (*4*) other candidates.

Figure 2 plots the indices of stellar photometric variability—the maximum amplitudes of the rotational modulation ΔV and of slow changes of the mean brightness $\Delta\langle V\rangle$—as functions of $V–I$. We do not detect any clear dependence of the photometric variability on spectral type (or $V–I$). However, it is clearly apparent from this figure that the strongest variability is reached for *K* stars ($V–I \approx 1\overset{m}{.}3$–$2\overset{m}{.}6$). The photometric variability is weaker for relatively red and blue sample stars. A marked decrease in the effects of starspots is observed for solar-type stars. A similar effect was noted by O'Dell *et al.* [20], who analyzed the relation between the rotational modulation amplitude ΔV and $B–V$ for young solar-type stars in the Pleiades, α Per cluster, and the general field. They suggested that the

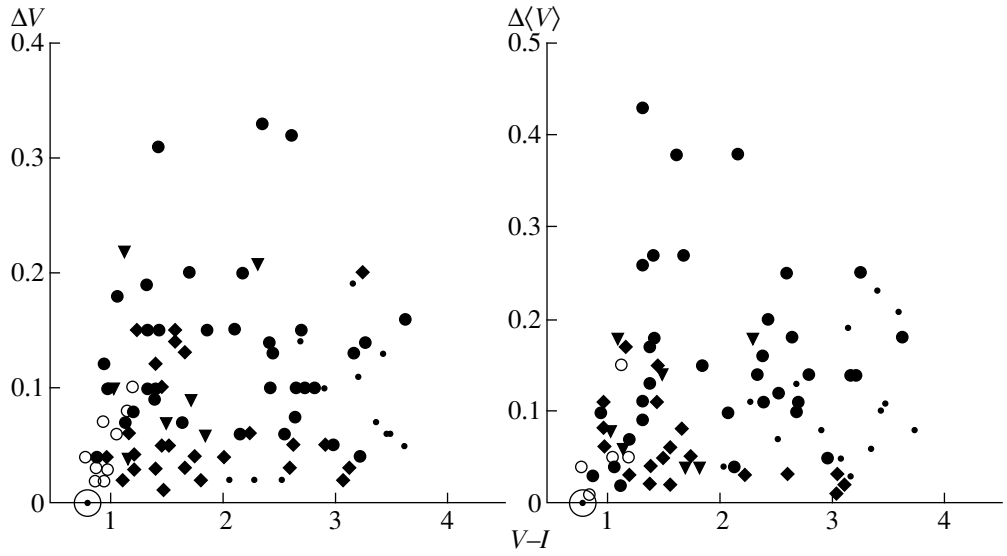


Fig. 2. Photometric variability of BY Dra stars as a function of *V–I*. Same notation as in Fig. 1.

lower variability displayed by hot stars was due to their smaller convective cells [21].

The rotational periods are known for 84 stars. The periods of most of these stars (65%) are rather short (below 5 days). The periods of 16% of the variables are between five and ten days; for 14%, they are between 10 and 20 days; and only 5% of all the stars have periods longer than 20 days. The longest period (41 days) was found for V645 Cen (Proxima Centauri). The corresponding rotational velocities are usually 10–20 km/s, but can be appreciably higher. For example, $V\sin i$ for TY Col is 55 km/s [22], for BI Cet and EY Dra, 60 km/s [23, 24], and for LO Peg and HK Aqr, 70 km/s [25, 26]. The rapid rotation of these stars provides evidence that they are young (for example, HK Aqr, LO Peg, EY Dra) or members of binary systems (BI Cet).

In addition to the known relation between a star's rotation and its age, there is also a relation between rotation and mass [27, 28]. A comparison of the rotational periods and *V–I* for spotted field stars shows no clear dependence between $P_{\rm rot}$ and *V–I*, since these stars have a variety of ages. The range of possible rotational periods becomes wider towards hotter stars. Three stars contradict this tendency: V1285 Aql (dM3e), V645 Cen (dM5e), and V1803 Cyg (dK5e), all of which show only weak variability.

The wider range of rotational periods towards hotter stars is due to the combined influence of the dependence of the rotation on mass and age. The "blue" side of the distribution shows both stars that have already reached the main sequence and stars that have not yet reached it, whereas almost all the cool stars have not yet reached the main sequence. The picture is further complicated by the short-term speedup in the rotation of stars with ages from 10^7 to 10^8 years, which can lead to different periods for stars with very similar colors and luminosities [28, 29].

Let us now consider the relation between BY Dra variability and a star's rotation. According to Bopp and Fekel [30] and Bopp and Espenak [31], spots show up at rotational velocities exceeding 5 km/s. However, this criterion is by no means strict, and slowly rotating stars showing spot activity—primarily via changes in their mean brightnesses—are known. In their study of the relationship between the rotational modulation amplitude and rotation for young solar-type stars, O'Dell *et al.* [20] noted a growth of ΔV with decreasing $P_{\rm rot}$. They suspected saturation at the period $P_{\rm rot} \sim 0\overset{\rm d}{.}5$.

Figure 3 displays our relations between the photometric variability of spotted field stars and the logarithm of their rotational periods. This figure shows that, for $P_{\rm rot} > 3^{\rm d}$ (or $V_{\rm rot} < 20$ km/s), the maximum observed amplitudes of the rotational modulation ΔV and of the mean brightness variations $\Delta\langle V\rangle$ generally increase towards more rapidly rotating stars. A similar relation was found by Granzer and Strassmeier [32] based on data for 87 active dwarfs.

The relation between photometric variability and period does not hold for VY Ari, which shows significant variability with a long ($16\overset{\rm d}{.}42$) rotational period. Note that this star is often classified as a K3IVe RS CVn system [11]. According to the statistical data [32], the rotational modulation amplitude is independent of period for active giants and subgiants.

The strongest photometric variability is observed for rotational velocities of 15–20 km/s and periods near three days (V775 Her, CC Eri). In such stars, the spotted regions can cover as much as 40 to 50% of the stellar surface [33]. The lower variability amplitudes for stars with higher rotational velocities are probably due to a saturation effect. This result is in contradiction with the results of O'Dell *et al.* [20], but in complete agree-

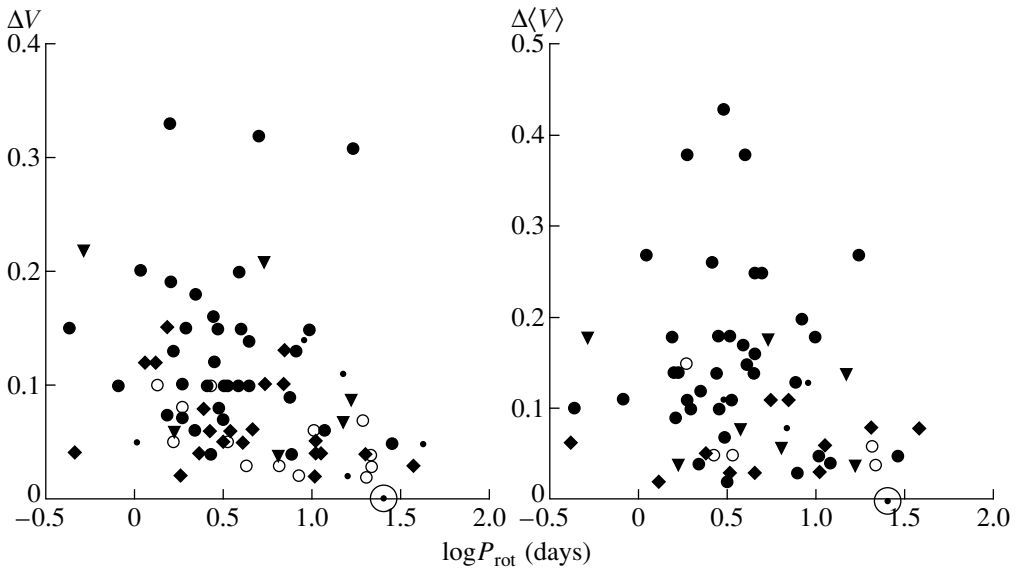


Fig. 3. Photometric variability of BY Dra stars as a function of rotational velocity. Same notation as in Fig. 1.

ment with similar dependences for other activity indicators [1–6].

One of the fundamental parameters of a cool star is its Rossby number, which depends on the star's spectral type and rotation [34]. Overall, the dependence of variability amplitude on Rossby number resembles that of Fig. 3: in general, the amount of possible photometric variability increases with decreasing $\log \mathrm{Ro}$. This provides evidence that the star's rotation directly influences its variability, while the effect of the convective-zone structure is not as strong. The dependences of ΔV and $\Delta\langle V\rangle$ on Ro also show signs of saturation at ~ 0.2–0.3, as is true of other activity indicators, whereas the sample of O'Dell *et al.* [20] does not show any evidence for saturation at such Ro values.

3. CONCLUSIONS

We have carried out a statistical survey of 113 spotted red dwarf stars with known or suspected BY Dra-type variability. The conclusions listed below follow from our comparison of quantities characterizing the variability (the rotational modulation amplitude ΔV and mean brightness variability amplitude $\Delta\langle V\rangle$) with the global parameters of the stars.

(1) The dependence of photometric variability on spectral type is weak, but there is some tendency for lower variability toward the blue and the red ends of the sample.

(2) Variability becomes stronger with increasing rotational velocity, possibly with saturation at the critical velocity $V_{\mathrm{crit}} \sim 15$–$20$ km/s.

(3) Photometric variability becomes stronger with decreasing Rossby number, also with saturation at $\mathrm{Ro}_{\mathrm{crit}} \sim 0.2$–$0.3$.

(4) The photometric variability of the Sun as a star fits in with all the above dependences.

Thus, we conclude that spot-forming activity on red dwarf stars depends on the star's mass, rotation, and Rossby number, as in the case of other manifestations of stellar activity.

REFERENCES

1. R. Pallavicini, L. Golub, R. Rosner, *et al.*, Astrophys. J. **248**, 279 (1981).
2. M. Mathioudakis, A. Fruscione, J. J. Drake, *et al.*, Astron. Astrophys. **300**, 775 (1995).
3. D. Montes, M. J. Fernández-Figueroa, E. de Castro, and M. Cornide, Astron. Astrophys. **294**, 165 (1995).
4. A. Scumanich, Astrophys. J. **171**, 565 (1972).
5. S. Catalano and E. Marilli, Astron. Astrophys. **121**, 190 (1983).
6. O. Vilhu and F. M. Walter, Astrophys. J. **321**, 958 (1987).
7. I. Yu. Alekseev and R. E. Gershberg, Astron. Zh. **73**, 589 (1996) [Astron. Rep. **40**, 538 (1996)].
8. B. R. Pettersen, Mem. Soc. Astron. Ital. **62**, 217 (1991).
9. R. E. Gershberg, M. M. Katsova, M. N. Lovkaya, *et al.*, Astron. Astrophys., Suppl. Ser. **139**, 555 (1999).
10. K. G. Strassmeier, D. S. Hall, M. Zeilik, *et al.*, Astron. Astrophys., Suppl. Ser. **72**, 291 (1988).
11. K. G. Strassmeier, D. S. Hall, F. C. Fekel, and M. Scheck, Astron. Astrophys., Suppl. Ser. **100**, 173 (1993).
12. G. W. Henry, F. C. Fekel, and D. S. Hall, Astron. J. **110**, 2926 (1995).
13. J. T. Hooten and D. S. Hall, Astrophys. J., Suppl. Ser. **74**, 225 (1990).
14. G. Cutispoto, R. Pallavicini, L. Pasquini, *et al.*, in *Cool Stars, Stellar Systems and the Sun*, Ed. by M. S. Giam-

papa and J. A. Bookbinder, Astron. Soc. Pac. Conf. Ser. **26**, 119 (1992).

15. G. Cutispoto, G. Tagliaferri, P. Giommi, *et al.*, Astron. Astrophys., Suppl. Ser. **87**, 233 (1991).
16. G. Cutispoto, G. Tagliaferri, R. Pallavicini, *et al.*, Astron. Astrophys., Suppl. Ser. **115**, 41 (1996).
17. D. S. Hall, in *The Sun and Cool Stars: Activity. Magnetism. Dynamos*, Ed. by I. Tuominen, D. Moss, and G. Rüdiger (Springer, Heidelberg, 1991), p. 353.
18. P. F. Chugaĭnov, Astrofizika **30**, 271 (1991).
19. N. I. Shakhovskaya, in *Flares and Flashes*, Ed. by J. Greiner, H. W. Dürbeck, and R. E. Gershberg, Lect. Notes Phys. **454**, 61 (1995).
20. M. A. O'Dell, P. Panagi, M. A. Hendry, and A. Collier Cameron, Astron. Astrophys. **294**, 715 (1995).
21. M. Bünte and S. H. Saar, Astron. Astrophys. **271**, 167 (1993).
22. G. Tagliaferri, G. Cutispoto, R. Pallavicini, *et al.*, Astron. Astrophys. **285**, 272 (1994).
23. B. W. Bopp, T. B. Ake, B. D. Goodrich, *et al.*, Astrophys. J. **297**, 691 (1985).
24. R. M. Robb and R. D. Cardinal, Inf. Bull. Var. Stars, No. 4270 (1995).
25. R. D. Jeffries, P. B. Byrne, J. G. Doyle, *et al.*, Mon. Not. R. Astron. Soc. **270**, 153 (1994).
26. B. R. Petterson, D. L. Lambert, J. Tomkin, *et al.*, Astron. Astrophys. **183**, 66 (1987).
27. G. W. Lockwood, D. T. Thompson, R. R. Radick, *et al.*, Publ. Astron. Soc. Pac. **96**, 714 (1984).
28. F. van Leeuwen and P. Alphenaar, ESO Messenger **28**, 15 (1982).
29. J. R. Stauffer, L. W. Hartmann, J. N. Burnham, and B. F. Jones, Astrophys. J. **289**, 247 (1985).
30. B. W. Bopp and F. C. Fekel, Astron. J. **82**, 490 (1977).
31. B. W. Bopp and F. Espenak, Astron. J. **82**, 916 (1977).
32. Th. Granzer and K. G. Strassmeier, in *Proceedings of the 11th Cambridge Workshop "Cool Stars, Stellar Systems and the Sun,"* Ed. by R. J. García López, R. Rebolo, and M. R. Zapatero Osorio, Astron. Soc. Pac. Conf. Ser. (in press).
33. I. Yu. Alekseev and R. E. Gershberg, Astron. Zh. **74**, 240 (1997) [Astron. Rep. **41**, 207 (1997)].
34. R. W. Noyes, L. W. Hartmann, S. L. Baliunas, *et al.*, Astrophys. J. **279**, 763 (1984).

Translated by N. Samus'

Astronomy Reports, Vol. 44, No. 10, 2000, pp. 701–710. Translated from Astronomicheskiĭ Zhurnal, Vol. 77, No. 10, 2000, pp. 790–800.

Evolution of Solar Magnetic Tubes from Observations of Stokes Parameters

V. A. Sheminova and A. S. Gadun

Main Astronomical Observatory, National Academy of Sciences of Ukraine, Goloseevo, Kiev, 252127 Ukraine

Received September 1, 1999

Abstract—Basic scenarios and mechanisms for the formation and decay of small-scale magnetic elements and their manifestation in synthesized Stokes profiles of the Fe I 15648.5 Å infrared line are considered in the context of two-dimensional modeling of nonstationary magnetogranulation on the Sun. The stage of convective collapse is characterized by large redshifts in the *V* profiles accompanied by complete Zeeman splitting of the *I* profiles. This is due to intense downward flows of material, which facilitates the concentration of longitudinal field with an amplitude of about several kG in the tube. The dissipation of strong magnetic structures is characterized by blueshifts in their profiles, which result from upward fluxes that decrease the magnetic field in the tube. Typical signatures during key stages in the evolution of compact magnetic elements should be detectable via observations with sufficiently high spatial and temporal resolution. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Studies of the interaction of thermal convection with magnetic fields via numerical modeling of magnetoconvection [1–7] are of considerable interest for our understanding of the structure and dynamics of small-scale magnetic fields, which carry most of the magnetic flux going from the solar photosphere, outside of pores and sunspots. Current time-dependent magnetohydrodynamic (MHD) models can successfully describe the observed concentration of magnetic flux in intergranular regions and explain variations in convection under the influence of magnetic fields. They also enable detailed study of mechanisms for the formation of compact magnetic elements and their evolution. Unfortunately, verification of MHD models and direct comparison with observations are not easy for a number of reasons, such as the limited spatial resolution of the observations.

One link between observations and MHD models is provided by Stokes diagnostics, which are now well developed. These represent a set of special methods for extracting information about the structure and dynamics of magnetic elements from spectropolarimetric observations. The use of Stokes diagnostics to study the structure of small-scale magnetic elements is necessitated by the fact that these elements are smaller (70–300 km) than the spatial resolution of modern instruments, so that the observed signal carries information not only about the magnetic structures, but also about the nonmagnetic surrounding environment. As a result, correct interpretation of such observations is difficult.

The aim of the present work is to study evolutionary changes in the structure and dynamics of magnetic elements, as well as variations in the parameters of synthesized Stokes profiles, via two-dimensional modeling of nonstationary magnetogranulation. We also wish to identify diagnostic signatures characteristic of magnetic-field intensification and dissipation.

The numerical modeling of nonstationary magnetoconvection we consider here was carried out for two hours of real solar time, and describes both the formation and dissipation of small-scale magnetic elements. In particular, precisely such modeling demonstrated for the first time the importance of the surface mechanism (scenario) for the formation of compact magnetic elements, which is realized during the fragmentation of large-scale thermal fluxes. This mechanism is described in detail by Gadun *et al.* [9], who also present the preliminary results of the modeling. These results are analyzed in more detail in [7], along with investigations of the formation and decay of tubes, tube stability regimes, differences in the brightness characteristics of magnetic and nonmagnetic granulation, and the dependences of magnetic-element parameters on their horizontal scales and the field intensity. Extensive test–diagnostic calculations—in particular, the synthesis of Stokes profiles of several spectral lines and comparison with observations—were conducted for the same models in [8]. These models adequately describe the basic features of spectropolarimetric observations of compact magnetic elements. Another important result of [8] was testing the available Stokes diagnostics methods, which was done by comparing the model parameters with those derived from an analysis of theoretical spectropolarimetric scans.

The present paper represents a continuation of these studies. It describes in more detail several critical stages in the evolution of compact magnetic elements that can be detected via spectropolarimetric methods. As usual, the small-scale magnetic elements in the numerical modeling will be called tubes, although they

1063-7729/00/4410-0701$20.00 © 2000 MAIK "Nauka/Interperiodica"

do not have a tube shape in a two-dimensional planar representation.

2. MHD MODELS AND CALCULATION OF STOKES PROFILES

We estimated all the thermodynamic parameters of the atmosphere required to solve the transfer equations for the spectral lines in the presence of a magnetic field via numerical MHD modeling of planar magnetoconvection of the granulation scales. We made the approximation that the medium was compressible and partially ionized, and was stratified by gravity and coupled with the radiation. The complete system of radiative MHD equations and its solution for the particular case at hand are described in [9]. As an initial model for the MHD simulations, we used nonmagnetic, two-dimensional models whose computation was based completely on the approaches presented in [10].

The upper and lower boundary conditions in the MHD simulations were taken to be free; i.e., there was free inflow and outflow of material. The velocity components were determined by the condition $\partial \mathbf{V}/\partial z = 0$, and the average values of the intrinsic energy and density were fixed by the initial uniform model [10, 11]. The profiles of the fluctuations of these quantities at the upper (or lower) boundary coincide with the corresponding variations for the layers located at lower (or greater) heights. In addition, the density at the lower boundary was scaled so that the sum of the gas, radiative, and magnetic pressures was constant at a fixed horizontal level.

The upper and lower boundary conditions for the magnetic field were specified by its global character and took the form $B_x = 0$ and $\partial B_z/\partial z = 0$. The lateral boundary conditions were taken to be periodic. The initial configuration of the magnetic field was bipolar, with the field intensity decreasing with height. The average value of B over the entire computational region was 54 G. We chose this configuration to help ensure numerical stability of the solution at the initial simulation time, since the velocity field and thermodynamic quantities, on the one hand, and the magnetic-field characteristics, on the other hand, are not self-consistent in the initial model.

The computational region was 3920×1820 km in size, with spatial steps of 35 km. The atmospheric layers covered about 700 km. It is obvious that we cannot study the intrinsic structure of fine magnetic configurations in detail with such a step size, but it is adequate for following trends in the evolution of magnetic configurations, and reduces the computer time required for the calculation.

To study the evolution of magnetic elements, we considered a sequence of MHD models from time 1.5 min at the beginning of the simulation (when the average magnetic field was about 54 G) to time 120 min (when the average magnetic field was ≈500 G). This sequence contains 94 two-dimensional models with time intervals of one minute and 52 models with intervals of 0.5 min. The computational region contains 112 vertical columns (rays). We computed four Stokes profiles for each vertical column for a particular model as for a plane-parallel atmosphere, after which we performed a spatial average of the resulting profiles. The equations for the transfer of polarized radiation in a magnetic field (often called the Unno–Rachkovskiĭ equations) form a system of four first-order differential equations. We used a fifth-order Runge–Kutta–Felberg method to numerically solve these transfer equations. The detailed computation algorithm, description of the software, and behavior of the Stokes profiles as functions of the state of the medium and atomic parameters are given in [12–14]. The only substantial restriction in the computation of the Stokes profiles is the assumption of local thermodynamic equilibrium.

To study temporal variations in the Stokes profiles, we chose the Fe I 15648 Å infrared line, which possesses very promising diagnostic capability [15]. Along with its high magnetic sensitivity, the important advantages of this line are its low temperature sensitivity and the fact that it is formed in very deep layers of the photosphere (below $\log\tau_5 = -1$). As shown in [8], this line is especially suitable for LTE analyses of theoretical MHD models, in which the temperature fluctuations in the upper layers of the photosphere can be artificially overestimated due to the use of a gray approximation for the radiative transfer.

3. GENERAL CHARACTERISTICS OF NONSTATIONARY MAGNETOGRANULATION

The results of the numerical simulations show that the character of the evolution of thermal fluxes significantly affects the development of magnetic configurations. The dissipation of granules and merging of intergranular gaps result in the disappearance of small-scale magnetic fields or the formation of cluster-like structures, whereas the fragmentation of granules in the presence of horizontal photospheric fields leads to the formation of new compact magnetic configurations.

Figure 1 shows that the entire simulation interval can be subdivided into three stages [7, 9]: an initial period (up to 20 min), which is primarily determined by the initial conditions; a transition period (from 20 to 35 min), when there is a mutual reconstruction of the magnetic field and thermal convection; and a final period of self-consistent evolution of the nonstationary convection and magnetic field. The first clear signatures of the presence of compact magnetic tubes with strong fields (of the order of several kG) appear after approximately 50 min of real solar time. Bright points form in the intergranular gaps, and the tubes begin to glow. The positions of the bright points (Fig. 1a) correspond to regions of strong magnetic field whose orientation is predominantly vertical

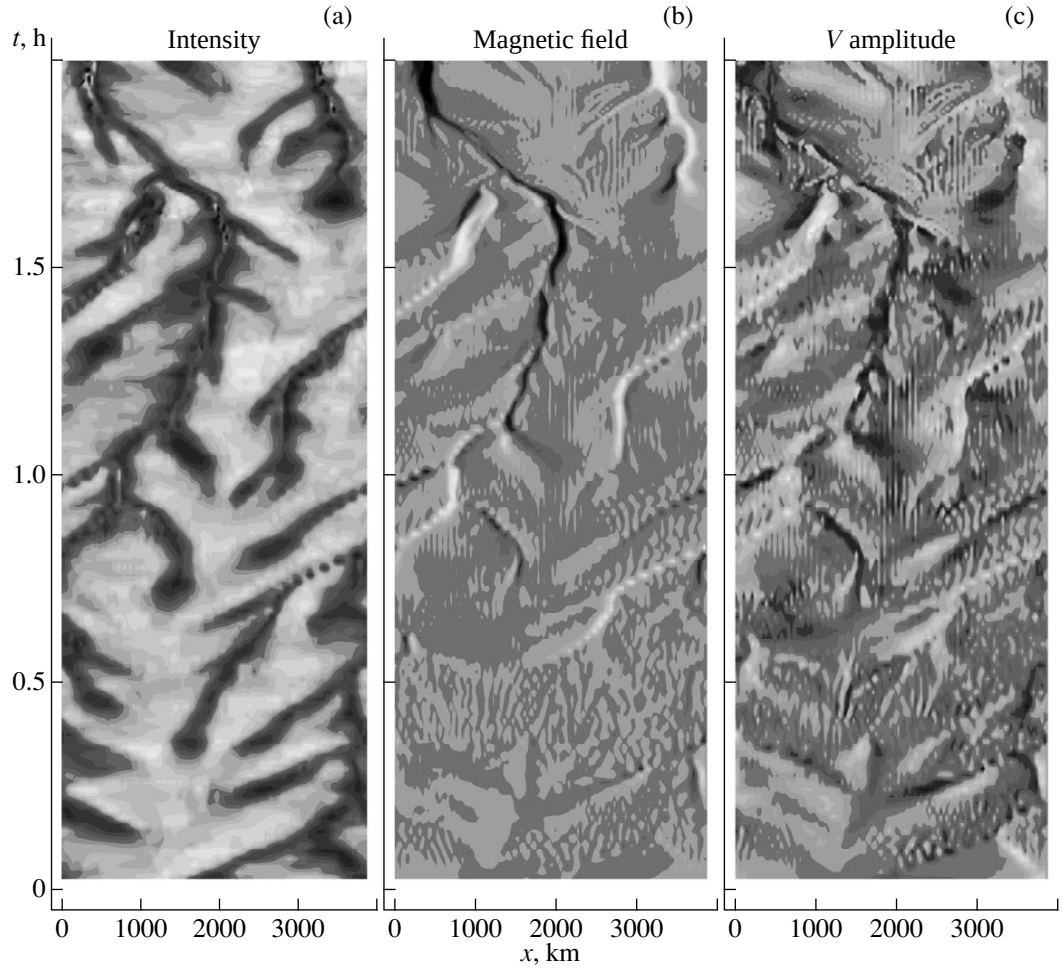


Fig. 1. Spatial and temporal evolution of (a) the intensity of monochromatic emission at wavelength 500 nm, (b) the magnetic-field intensity, and (c) the amplitude of blue wing of the Stokes *V* profile (b) in the course of two-hour, two-dimensional MHD modeling. The amplitude of the variations in the relative intensity in (a) is 0.017–1.78, and is represented by a linear gradation from dark to light shading. The amplitude of the variations in *B* in (b) ranges from –2867 to +2540 G. Negative and positive polarity are shown by dark and light shading, respectively. This intensity was determined for the $\log\tau_R = 0$ level, where the continuum is formed. The variations in the amplitude of the *V* profiles of the Fe I 15648 Å line in (c) range from –0.14 to 0.16. Negative and positive values are shown by dark and light shading, respectively.

or close to vertical (Fig. 1b), and also to regions of increase in temperature, decrease in pressure, and oscillating upward–downward flows in the intergranular gaps.

Due to the effect of the hot walls, the glow of the tubes will begin earlier near the disk edges than at their centers. This suggests the possibility of early observational diagnostics of magnetic elements that have just begun to be formed [3]. Note also that the compact structures with strong magnetic fields that form are often located near regions with relatively weak fields of the opposite direction. The character of the granulation also changes. The magnetic field produces a stabilizing action: there are no strong horizontal shears in the presence of the magnetic field, as distinct from the case of nonmagnetic granulation. The size distribution of the granules in the magnetogranulation also differs from the nonmagnetic case: the granules are smaller, and there are many small-scale bright structures with sizes less than 150–300 km associated with bright points of the magnetic tubes [7].

There are two scenarios for magnetic-tube formation in the numerical simulations under consideration [7]. The first is based on the concentration of a magnetic field between cells and its intensification up to the equipartition level due to a kinematic mechanism; the field can then be further amplified by the development of superconvective instability. The second mechanism is based on the formation of a tube during the decay of thermal convective flux. The magnetic-tube behavior in this case also exhibits two stages. The first takes place when the field intensity has not reached the equipartition level, so that the tube is in a state of convective instability. The subsequent, stronger stage occurs when the upper part of the tube is in radiative equilibrium and experiences an

oscillatory instability. The transition from one regime to the other occurs at $B \approx 1200$–1300 G [7].

The situation changes abruptly when convective collapse is "turned on." The level of the observable surface of the magnetic tube becomes more strongly dependent on B. The Wilson depression increases. The temperature at the corresponding level also increases, since the observable surface shifts to deeper layers, and the plasma configuration is actively heated by neighboring hotter thermal fluxes. The gas pressure decreases due to the increasing magnetic pressure, and the density also decreases. The tubes begin to glow, and the intensity of their emission is above the average level. The tubes are also surrounded by dark regions, which correspond to channels of intense downward flows of material near the tubes. The brightness peaks of most wide tubes are split due to the effect of the large horizontal scale, which results in less efficient lateral heating. As a result, the total contrast of the tubes increases.

Our simulations revealed three basic mechanisms for the disintegration of the tubes [7]: dissipation of the magnetic field due to reconnection of the field lines, disintegration due to interchange instability, and reversal of the convective collapse. The first case—dissipation of the magnetic field—is the most common scenario. It occurs when the thermal flux separating two tubes with opposite field directions disappears, and the tubes merge. Interchange instability takes place when the orientation of intense magnetic tubes deviates from the vertical. The rising of strong horizontal magnetic field results in the development of flute instability. In the plane case under consideration, this instability can split and dissipate strong compact tubes. Inverse convective collapse [6] is produced by the strong depletion of the upper part of the tube when it is intensified. Such depletion reverses the convective collapse, so that downward flows in the tube are replaced by upward ones, whose velocity can reach supersonic values in the atmosphere [6].

The evolution of magnetoconvection briefly outlined above is analyzed in more detail in [7]. Here, we will compare separate critical stages of tube evolution with Stokes profiles synthesized for the same simulation times. With this aim in view, we chose the most typical cases for the formation and decay of magnetic tubes from the temporal sequence of MHD models.

4. EXAMPLES OF EVOLUTION OF A KILOGAUSS MAGNETIC TUBE

Figure 1c shows variations in the amplitudes of the blue wings of the synthesized Stokes V profiles over two hours in the simulation region. These changes are very similar to the space–time variations in the magnetic field (Fig. 1b). To illustrate the formation and dissipation of the magnetic tubes, we shall consider in detail only two separate regions of the MHD models, whose evolution is presented in Figs. 2 and 3. The thick curve in the temperature distribution denotes the level $\tau_R = 1$ (where τ_R is the Rosseland opacity), and the dotted curve corresponds to the lowest temperature (4000 K). The vertical dashed line indicates the central axis of the region in which the magnetic tubes are located. Figures 4 and 5 present Stokes profiles of the Fe I 15648 Å spectral line with various spatial resolutions with respect to the central tube axis—35, 105, 175, and 315 km—synthesized for these regions. Note that a resolution of 35 km corresponds to the horizontal step of the simulation, so that the corresponding profiles were calculated at the tube center without spatial averaging; they are marked by the solid curves in the figures. We determined the parameters commonly used in Stokes diagnostics precisely for these profiles. These parameters are the amplitude of the V profile, $a_V = (a_b + a_r)/2$ (where b and r are the blue and red wings), the magnetic-field intensity (in G) derived from the distance between the maximum peaks of the V profile, $B_{br} = (\lambda_r - \lambda_b)/(2 \times 4.67 \times 10^{-13} \lambda_0^2 g_{\text{eff}})$ (where λ_0 is the wavelength of the unshifted line center in Å), the shift of the V profile, $V_V = c(\lambda_z - \lambda_0)/\lambda_0 - 2.12 \times 10^{-6}$ s (where λ_z is the wavelength of the zero intersection of the profile), and the relative asymmetry of the amplitudes and areas of the V profile, $\delta a = (a_b - a_r)/(a_b + a_r)$ and $\delta A = (A_b - A_r)/(A_b + A_r)$.

4.1. Formation of a Magnetic Tube

Let us compare the calculated Stokes profiles and their parameters (Fig. 4) with the physical processes of magnetic-tube formation due to the surface mechanism (Fig. 2).

A situation typical for the beginning of fragmentation of a large convective cell with a diameter of about 3000 km occurs in the (3920×1820 km) computational region at time 57 min. There is an upward convective flow in its central part (x = 2100 km). Two reverse flows, directed from the central regions, can be seen at the periphery. A spot of high-density material appears in the upper layers of the photosphere. The temperature at the $\log \tau_R = 0$ level decreases in regions of weak flow due to radiative cooling. There is a weak horizontal magnetic field with varying direction in the cell. In the predicted region of magnetic-tube birth (whose center is marked by the vertical dashed line in Fig. 2), the calculated Stokes profiles are shifted toward the blue part of the spectrum, $V_V = -0.8$ km/s. The minus sign denotes upward motion. The weak V profile (drawn by the solid curve in Fig. 4) includes two components with opposite polarities (the negative one is quite strong, and the positive one very weak).

The situation at time 59 min is typical for the initial phase of magnetic-tube birth. A downward jet of cold material forms during the fragmentation of the thermal flux (a convective cell). This flow drags the photospheric magnetic field with it, so that the horizontal

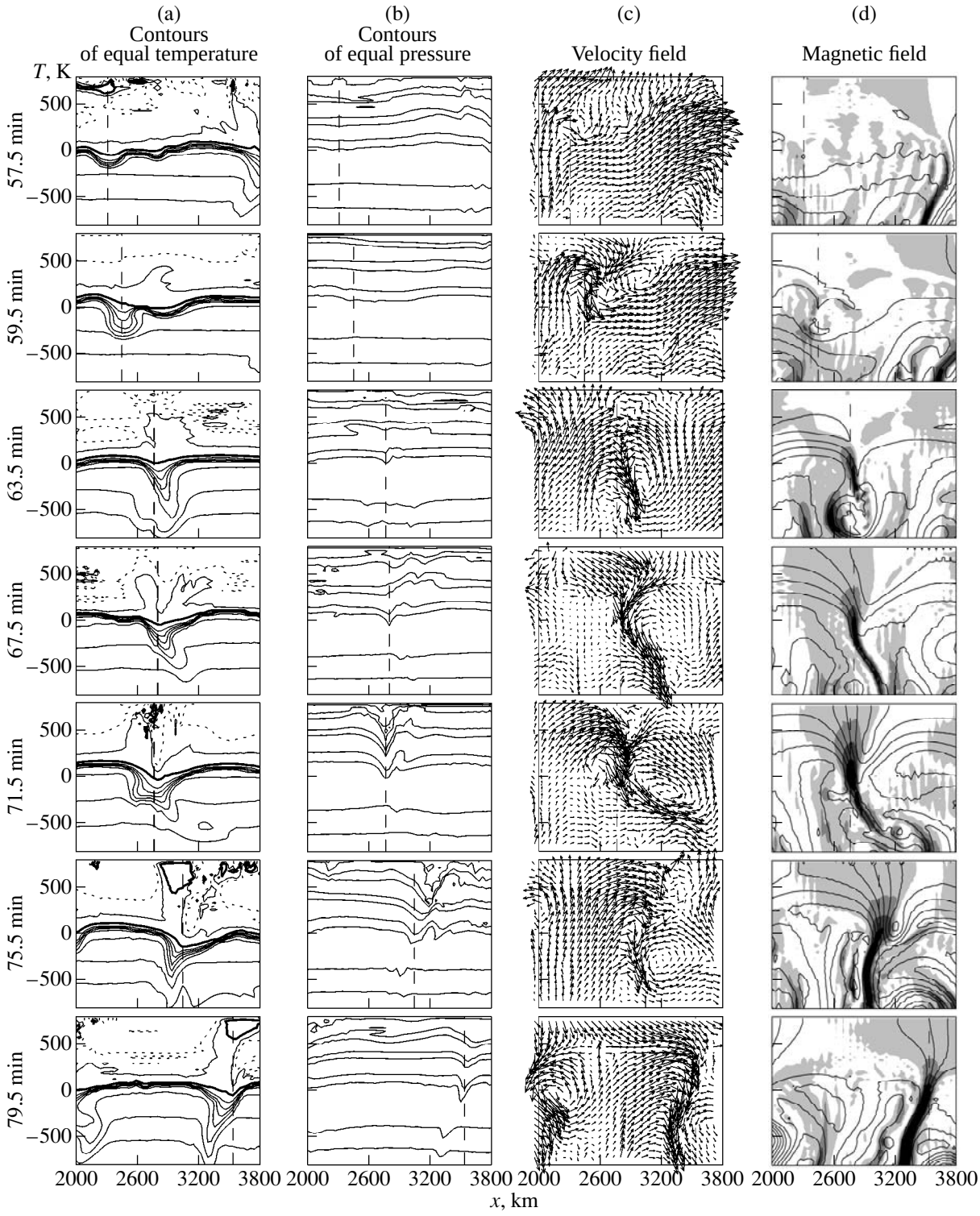


Fig. 2. Spatial and temporal evolution of magnetic-tube formation for the part of the simulated region from 2000–3800 km in the time interval 57–79 min: distribution of the (a) temperature T (the dotted curve corresponds to 4000 K and the thick curve to 6000 K, i.e., to the $\tau_R = 1$ level), (b) gas pressure P, (c) velocity field V, and (d) magnetic-field intensity B (the density of the shading is proportional to the values 10, 400, 800, and 1200 G; dark and light regions correspond to positive and negative polarity, respectively). The vertical dashed line denotes the center localization of the tube.

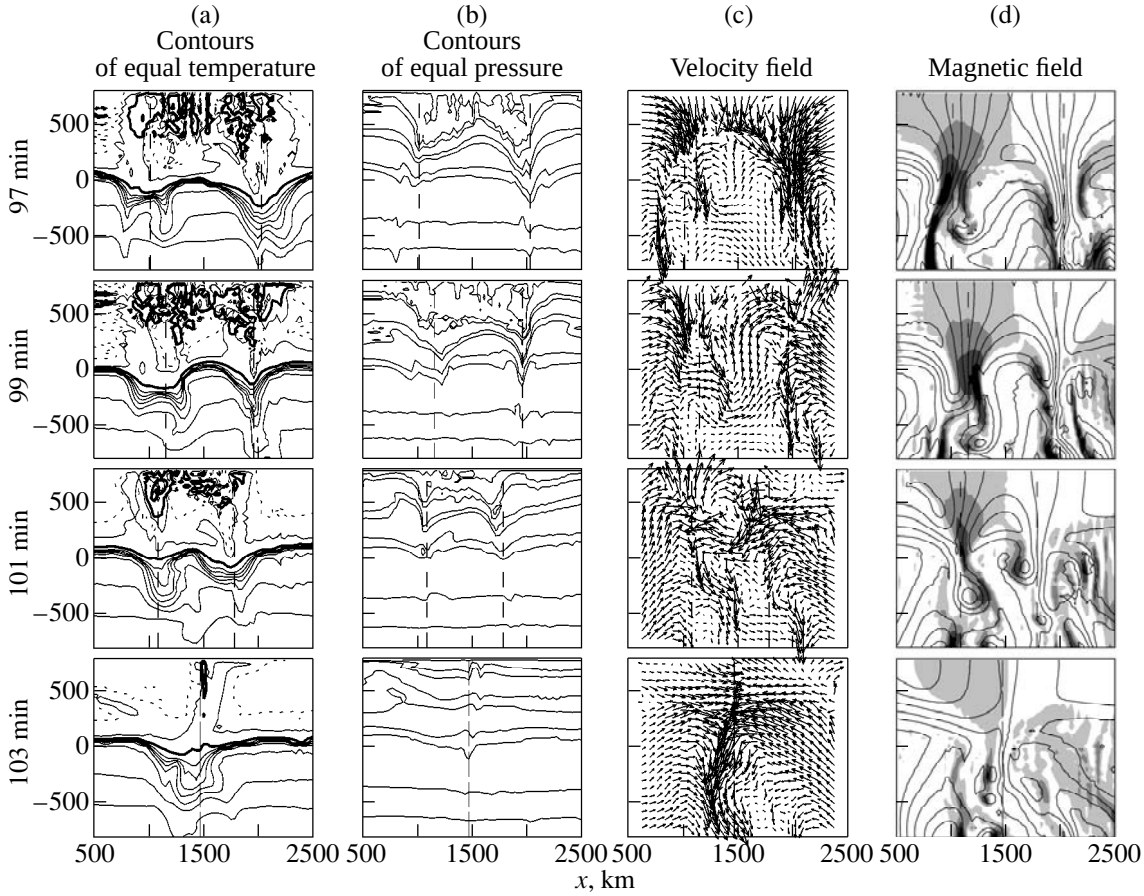


Fig. 3. Spatial and temporal evolution of the dissipation of two magnetic tubes for the part of the simulated region from 500–2500 km. All notation is the same as in Fig. 2.

field is transformed into a longitudinal field. The *I*, *Q*, and *V* Stokes profiles are shifted toward the red part of the spectrum, and V_V = 3.0 km/s. The changes in the shape of the *V* profile (where the positive component is dominant) are the result of the increase in the longitudinal field with positive polarity.

Further, the formation of the downward jet continues at times 63–67 min. The stratification of the material increases. There is a deficit of the gas pressure, and the positive magnetic field is sharply amplified. The accelerated vertical flows carry cold material and the magnetic field downward. Hot walls are formed in the upper part of the channel due to compression. The strong magnetic flux reaches the bottom of the computational region. The amplitude of the *V* profile takes on its maximum values. The complete splitting of the profiles indicates that the field intensity in the tube reached kilogauss values (B_{br} = 1078 G). The redshift of the profiles is 3.3 km/s.

The subsequent times 71, 75, and 79 min demonstrate the evolution of a fully formed vertical magnetic tube with diameter 150 km at the $\log\tau_R = 0$ level, which is in a state of convective collapse. The kilogauss longitudinal magnetic field in the channel enters deeper layers. The *I* Stokes profiles are completely split during this interval (B_{br} = 1246, 1195, and 1231 G), the wing amplitudes have decreased, and the redshifts are large—V_V = 4.9, 2.2, and 3.4 km/s.

Thus, this fragment of magnetoconvection considered over an interval of 22 min (from 57 to 79 min with respect to the beginning of the simulation) shows that the surface mechanism for magnetic-tube formation reaches the stage of convective collapse during its development. The downward flow of material amplifies the magnetic flux in the channel to kilogauss values within 6–10 min. In the next 12 min, convective processes around the tube gather and push the magnetic field lines toward the channel, leading to further concentration of the field in the magnetic tube. These processes have an oscillatory character, due to the presence of both global oscillations of the entire computational region and oscillations of the tube by itself. In general, the characteristic spectral features of the evolution revealed by the simulations confirm the results obtained in [6].

4.2. Decay of Magnetic Tubes

Let us now consider a scenario for magnetic-field dissipation due to the forced merging of two tubes with opposite field directions, which is most typical of the

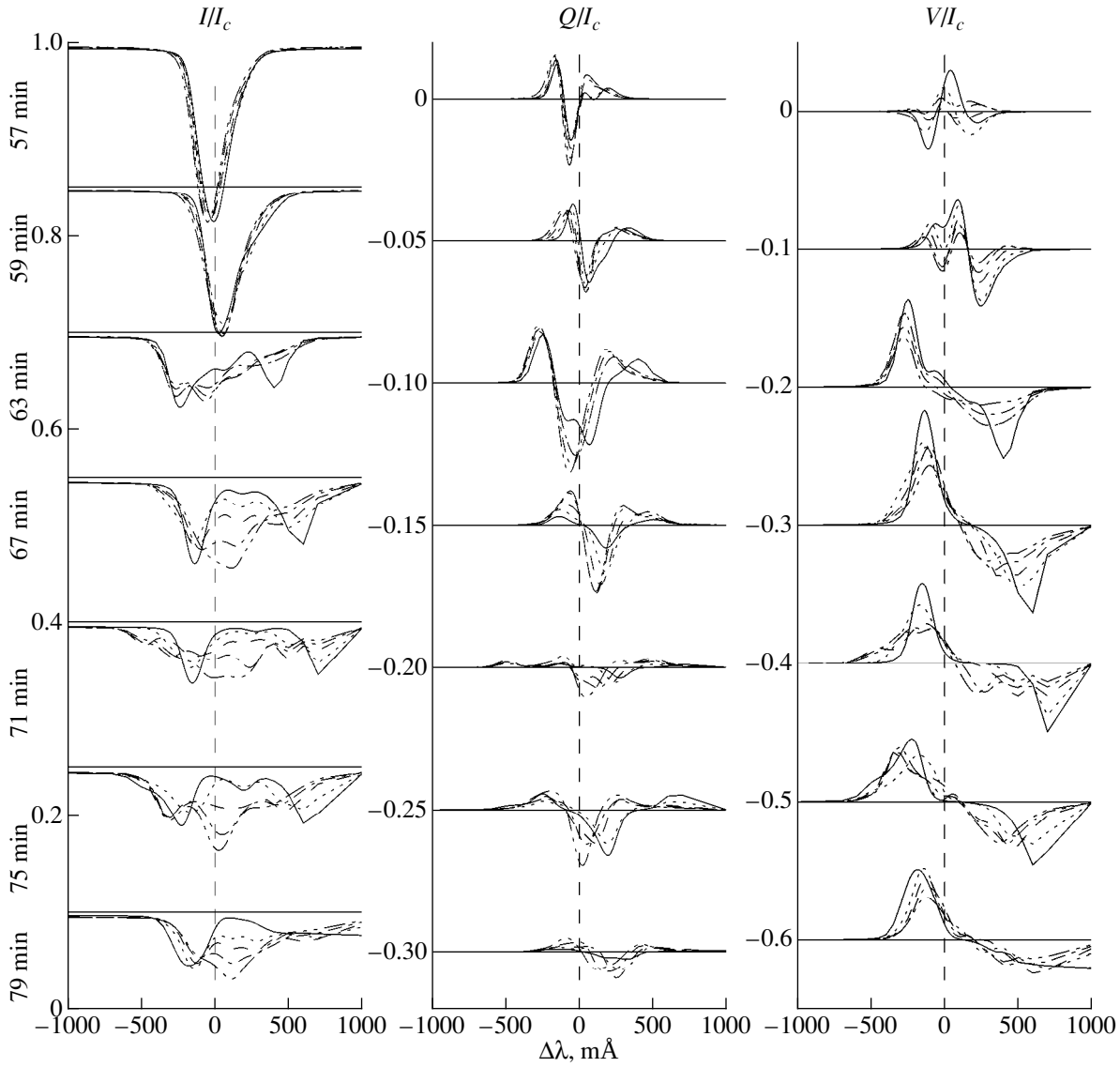


Fig. 4. Stokes profiles of the Fe I 15648 Å infrared line calculated at the center of the region where the kilogauss tube is formed for various spatial averaging scales: 35 km (solid curve), 105 km (dotted curve), 175 km (dashed curve), and 315 km (dot-dash curve).

simulations considered. The dynamics of these processes are shown in detail in Fig. 3, and the responses of the Stokes profiles calculated for each tube separately are presented by the solid curve in Fig. 5.

We can see two kilogauss vertical tubes with different field directions at time 97 min, separated by a distance of about 1000 km. The disintegration begins in the first, older, tube, which has a diameter of 350 km and positive polarity: the concentration of the magnetic flux decreases due to the tube's separation and the appearance of upward motions in deep layers inside the tube. The magnetic field has already reached the bottom of the computational box; i.e., the kilogauss longitudinal magnetic field has entered deep layers. The profiles are completely split (B_{br} = 1384 G), have a regular shape, and are redshifted by 1.1 km/s.

The second tube, which has a smaller diameter (250 km) and negative polarity, is a well-developed intense magnetic tube with some signatures of superconvective instability: there are accelerated downward flows of material, a maximum concentration of longitudinal magnetic field along the entire tube down to the base of the model, a narrow region of downward motion with a gas-pressure deficit, a rarefied region at the top, and a strong Wilson depression. At this time, the distance between the peaks of the calculated V profile (B_{br} = 1530 G) and the redshift (V_V = 5.0 km/s) are maximum.

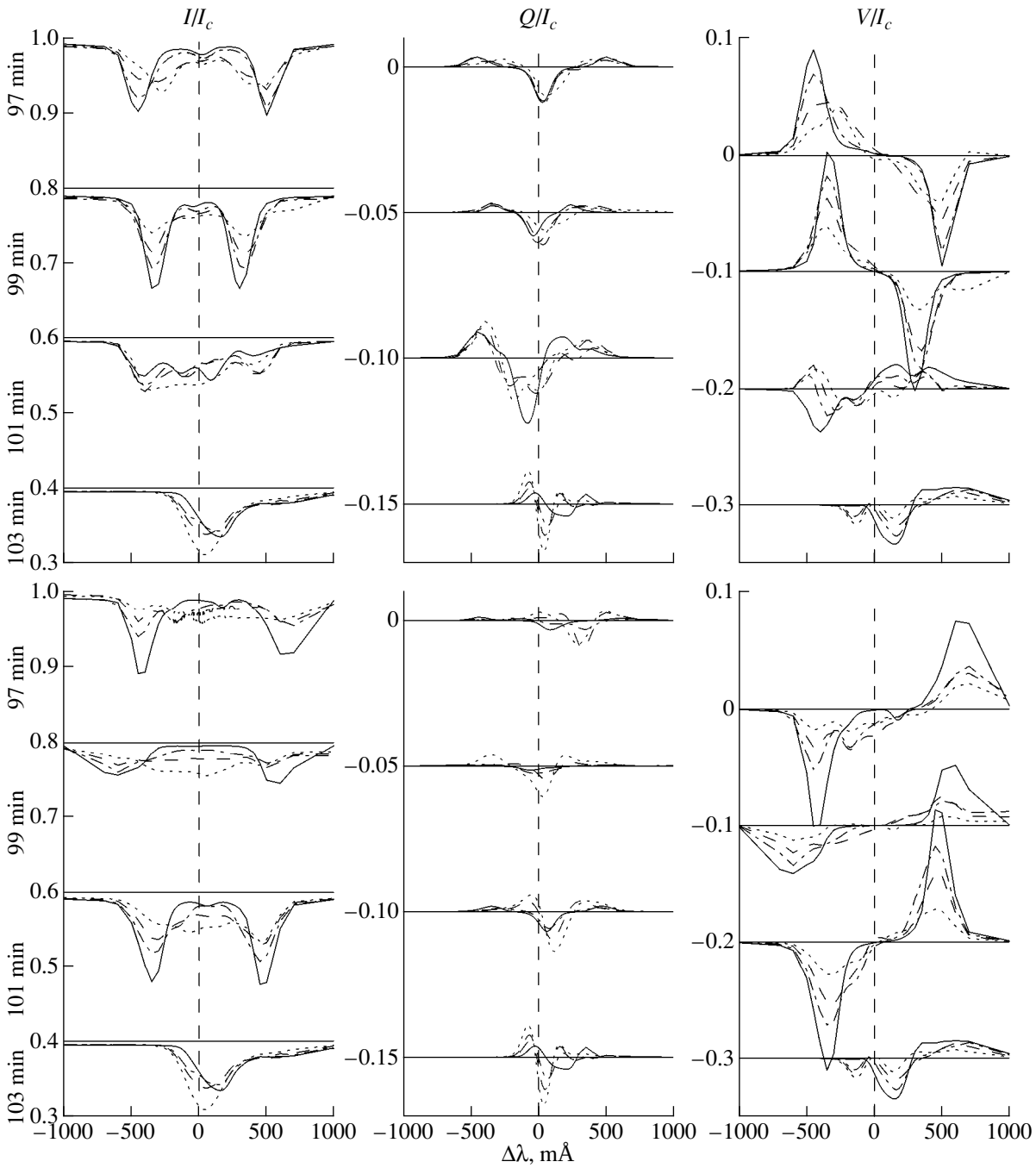


Fig. 5. Stokes profiles of the Fe I 15648 Å line calculated at the centers of regions where there is dissipation of a tube with positive magnetic-field polarity (top panels) and another tube with negative polarity (bottom panels). All notation is the same as in Fig. 4.

At subsequent times 99–101 min, the thermal flux separating the tubes disappears, and the two tubes with opposite magnetic-field directions begin to move toward each other, resulting in reconnection of the field lines. Characteristic motions similar to siphon flow can be seen between the tubes, whose features are described in detail in [16]. In the case under consideration, only magnetic-field lines between the tubes (rather than in the entire tube volume) participate in the siphon flows. This process decreases the magnetic field. In general, the dynamics of the motion are very complex. The flows of material in the tubes have considerably different velocity gradients and directions of motion. The upward flow increases in the positive tube, while the downward flow dominates inside the negative tube. Vortex motions appear between the tubes. All these processes result in disintegration of the tube structure and, accordingly, a rapid decrease in the magnetic-field concentration.

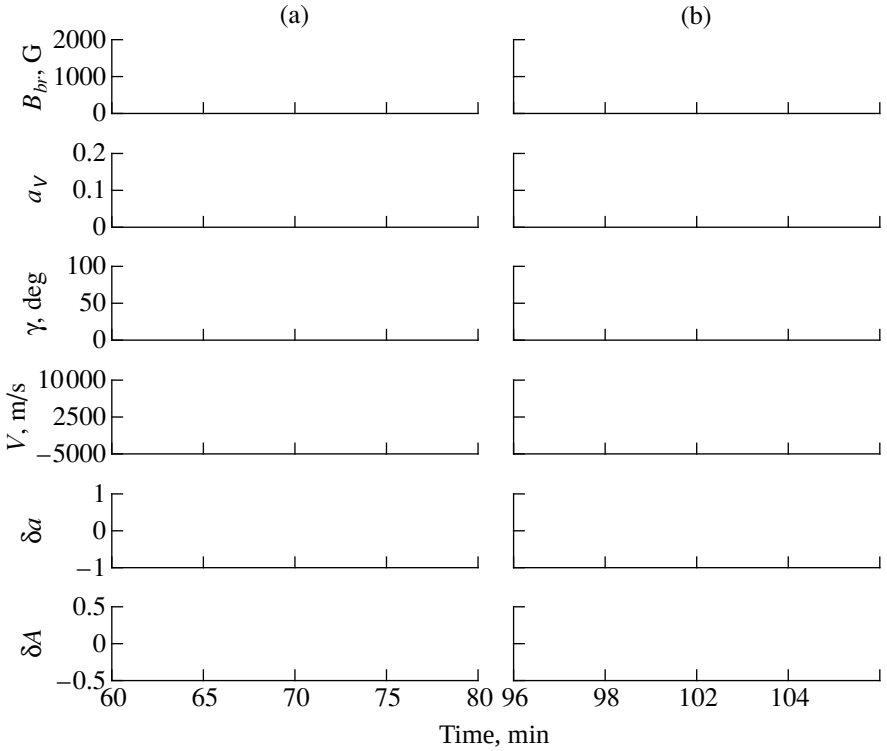


Fig. 6. Variations in the parameters of Stokes diagnostics during (a) formation of the magnetic tube and (b) dissipation of two magnetic tubes. The solid and dotted lines correspond to the tubes with positive and negative magnetic field polarity, respectively. Here, B_{br} is the magnetic-field intensity, a_V the amplitude of the V profile, V_V the vertical velocity, and δa and δA the asymmetries of the amplitudes and areas of the V profile.

At time 101 min, the calculated V profile gives V_V = –1.0 km/s, B_{br} = 700 G for the first tube and V_V = 0.7 km/s, B_{br} = 1166 G for the second tube. At time 103 min, there is only a weak field with negative polarity in place of the two kilogauss tubes. A new powerful flow of material begins to move downward along the existing thermal channel with its pressure deficit. The profiles are appreciably shifted toward the red part of the spectrum, and the right wing is very broad (V_V = 5.4 km/s, B_{br} = 656 G).

Therefore, within a very short time interval of 6 min (from 97 to 103 min with respect to the beginning of the simulation), dissipation of the thermal flux (convective cell) separating two tubes with opposite field directions led to merging of these compact magnetic elements and their disintegration.

4.3. Evolution of Magnetic Tubes and Variations in the Stokes Profiles

To trace the dynamics of temporal variations in the calculated Stokes profiles in the regions of formation and decay of the magnetic tubes, first and foremost, we studied the influence of spatial averaging on the results of Stokes diagnostics. This problem is considered in more detail in [8]. Figures 4 and 5 present the Stokes profiles calculated with various spatial averaging scales. A comparison of these profiles clearly shows that their shape substantially depends on the averaging region. The main reason for this is the complex structure of the simulated magnetic tube. The presence of appreciable horizontal and vertical gradients of the magnetic-tube parameters and of different directions for the magnetic field in the averaging region result in an anomalous shape for the Stokes profiles. The shapes are undistorted only for profiles calculated without horizontal averaging, i.e., for a single column in the tube center. Thus, only such profiles are able to reveal evolutionary spectral features. We used these profiles to calculate the parameters of the Stokes diagnostics, which are presented in Fig. 6 as functions of time.

As can be seen from the dependences in the left-hand side of Fig. 6, the following features are characteristic of the stage of magnetic-tube formation.

(1) The distance between the $\sigma_\pm$ components of the V profiles in units of the magnetic-field intensity B_{br} reaches a maximum at times 71–79 min, when the tube experiences convective collapse.

(2) The amplitude of the V profiles, a_V, increases, reaching its maximum just before the convective collapse, and then begins to decrease.

(3) Redshifts of the profiles V_V predominate over the entire formation period and lifetime of the magnetic tube. They are plotted in velocity units in Fig. 6. The

redshifts reach considerable values (up to 5 km/s) during the convective collapse.

(4) The amplitude asymmetry of the *V* profiles δa is positive, varies in the range 0.02–0.3, and is correlated with redshift.

(5) For the most part, the area asymmetry δA is negative and considerably less than δa.

The right-hand side of Fig. 6 presents the parameters of the Stokes diagnostics for the tube decay. The positive tube (solid line) begins to disappear at time 97 min, and the negative one (dotted line) at time 99 min. The resulting dependences can be summarized as follows.

(1) The magnetic splitting of the profiles decreases.

(2) The amplitude of the *V* profiles initially increases and then sharply decreases.

(3) Blueshifts are dominant at the initial stage of disintegration of the tube. They also can reach considerable values (up to 5 km/s), but are observed for a substantially shorter time than are large redshifts.

(4) The asymmetry of the *V*-profile amplitudes at the beginning of the disintegration is close to zero or is negative, but later becomes positive and increases.

(5) The asymmetry of the *V*-profile areas slightly decreases during the disintegration of the tube.

The time dependences for the Stokes-profile parameters lead us to recommend measurement of shifts of the zero intersection of the Stokes *V* profiles as a practical diagnostic tool. Such shifts can be reliably determined from observations with high spatial and temporal resolution.

5. CONCLUSION

We have considered typical cases for the formation and decay of kilogauss magnetic tubes in association with corresponding temporal variations in the Stokes profiles of the Fe I 15648 Å infrared line. The most characteristic signature of convective collapse amplifying the magnetic field inside tubes to kilogauss values is extremely large redshifts of the Stokes *V* profiles and completely split *I* profiles. On the contrary, the process of tube dissipation is characterized by blueshifts in the Stokes profiles, with typical features of weakening fields.

ACKNOWLEDGMENTS

We are very grateful to S.K. Solanki for detailed discussions of the results of this paper, valuable advice, and comments, and also to S.R.O. Ploner for various assistance. This work was supported by the Swiss National Scientific Foundation (grant no. 7UKPJ048440).

REFERENCES

1. Å. Nordlund, in *Proceedings of the Workshop on Small Magnetic Flux Concentrations in the Solar Photosphere*, Ed. by W. Deinzer, M. Knölker, and H. H. Voigt (Vandenhoeck and Ruprecht, Göttingen, 1986), p. 83.
2. O. Steiner, U. Grossmann-Doerth, M. Schüssler, and M. Knölker, Sol. Phys. **164**, 223 (1996).
3. P. N. Brandt and A. S. Gadun, Kinematika Fiz. Nebesnykh Tel **11** (4), 44 (1995).
4. Å. Nordlung and R. F. Stein, in *Solar Photosphere: Structure, Convection and Magnetic Fields (IAU Symposium no. 138)*, Ed. by J. O. Stenflo (Kluwer, Dordrecht, 1990), p. 191.
5. I. N. Atroshchenko and V. A. Sheminova, Kinematika Fiz. Nebesnykh Tel **12** (4), 32 (1996).
6. U. Grossmann-Doerth, M. Schüssler, and O. Steiner, Astron. Astrophys. **337**, 928 (1998).
7. A. S. Gadun, Kinematika Fiz. Nebesnykh Tel. **16** (2), 99 (2000).
8. V. A. Sheminova, Kinematika Fiz. Nebesnykh Tel **15** (5), 398 (1999).
9. A. S. Gadun, V. A. Sheminova, and S. K. Solanki, Kinematika Fiz. Nebesnykh Tel **15** (5), 387 (1999).
10. A. S. Gadun, Kinematika Fiz. Nebesnykh Tel **11** (3), 54 (1995).
11. A. S. Gadun and Yu. Yu. Vorob'ev, Astron. Zh. **73**, 623 (1996) [Astron. Rep. **40**, 569 (1996)].
12. E. Landi Degl'Innocenti, Astron. Astrophys., Suppl. Ser. **25**, 379 (1976).
13. V. A. Sheminova, Available from VINITI No. 2940-V90 (Kiev, 1990).
14. V. A. Sheminova, Preprint No. ITF-90-87 R (Institute of Theoretical Physics, Academy of Science of USSR, Kiev, 1991).
15. S. K. Solanki, I. Rüedi, and W. Livingston, Astron. Astrophys. **263**, 312 (1992).
16. J. H. Thomas and B. Montesinos, Astrophys. J. **375**, 404 (1991).

Translated by Yu. Dumin