Astronomy Reports, Vol. 49, No. 10, 2005, pp. 771–776. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 867–873.

Effective Curvature of the Universe in Observations of Distant Objects

E. V. Ivanova and O. S. Khovanskaya

Moscow State University, Leninskie gory, Moscow, 119899 Russia
Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia

Received November 28, 2004; in final form, May 18, 2005

Abstract—We analyze the propagation of light in a universe with density inhomogeneities, which, as was first demonstrated by Zel'dovich, systematically affect the angular sizes and apparent magnitudes of distant objects. The Zel'dovich effect can be represented in terms of an effective curvature, which is proportional to the magnitude of the inhomogeneities. We estimate the effective curvature, radius of curvature, and mean distance between conjugate points for the cases when the density inhomogeneities are taken to be stars in the Galaxy, galaxies, and clusters of galaxies. We discuss the possibility of detecting the Zel'dovich effect using modern astronomical data. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Small-scale deviations of the Universe from uniformity and isotropy due to individual objects such as stars, galaxies, and clusters of galaxies produce systematic distortions in cosmological tests [1, 2]. Zel'dovich [1] suggested corresponding correction terms, whose meaning is that a Universe whose spatial cross section has zero average curvature ($K = 0$) appears to some extent similar to an open cosmological model ($K < 0$). However, Zel'dovich represents the effect in the form of corrections to observational tests (the angular sizes and apparent magnitudes of distant objects), which cannot be conveniently transformed into geometrical parameters of the Universe, such as the curvature of the comoving space.

The modern physics of stochastic media can be used to describe the Zel'dovich effect in the following compact form [3]. The propagation of light in a Universe that is uniform only on average can be described by the Jacobi equation

$$y'' + K(x)y = 0 \tag{1}$$

with the initial conditions $y(0) = 0$ and $y'(0) = 1$. Here, y is the length of the geodetic deviation $\mathbf{y}$; x is the distance along the spacelike geodetic curve Γ; and $K = \bar{K} + k(x)$ is the curvature in the two-dimensional direction, which is the sum of some mean curvature $\bar{K}$ and random curvature variations $k(x)$ associated with deviations of the Universe from uniformity and isotropy. The behavior of the mean Jacobi field $\mathcal{Y} = \langle y \rangle$ is described by the equation

$$\mathcal{Y}'' + \mathcal{Y}K_{\text{eff}} = 0, \tag{2}$$

where

$$K_{\text{eff}} = \bar{K} - k^2\delta^2/3, \tag{3}$$

and $k = \sqrt{\langle k(x)^2 \rangle}$ is the rms fluctuation of the curvature and δ is the correlation scale length. Thus, an observer perceives a Universe with the critical density (and hence zero curvature) with small-scale deviations from uniformity and isotropy as a Universe with a negative effective curvature K_{eff} that is proportional to the amplitude of the inhomogeneities.

The second effect consists of the following. Let $\bar{K} = 0$ and the density fluctuations, and, consequently, the perturbations of the curvature k, be small. According to (3), the Universe is then characterized by a negative effective curvature. However, there is a nonzero probability that the set of areas of positive curvature is large enough to act as a gravitational lens; i.e., in geometric terms, to produce conjugate points. This effect is a result of the cumulative action of a large number of randomly distributed perturbations and supplements the microlensing effect [4–8], which produces conjugate points as a result of the deflection of light by an individual massive object acting as a gravitational lens. The mean distance between conjugate points of an object in whose light cone the inhomogeneities are located can be estimated as [9]

$$\Delta l \leq 4\delta/\nu, \tag{4}$$

where Δl is the distance between the conjugate points in which we are interested, δ is the mean distance between inhomogeneities, and ν is the deflection angle from a rectilinear trajectory of a light beam passing near an individual inhomogeneity.

Our aim is to relate the estimates (3) and (4) to observational data. We assume that the mean

1063-7729/05/4910-0771$26.00

curvature of space is $\bar{K} = 0$. As possible density inhomogeneities producing curvature fluctuations, we consider stars in our Galaxy, galaxies, and clusters of galaxies.

2. EFFECTIVE CURVATURE AND CURVATURE RADIUS

2.1. Stars

Let us assume that the Galaxy is uniformly filled with solar-mass stars ($m_{\rm star} \sim M_\odot$) with radii $r_0 \sim 7 \times 10^{10}$ cm separated by an average distance of $\delta \sim 4.6$ pc.

Using the Schwartzschild metric [10] to describe the gravitational field of the inhomogeneities, we obtain for the curvature perturbation $k(r)$ produced by a star at a distance r

$$k(r) = r_{\rm g}/r^3, \qquad (5)$$

where r is the distance from the field center to the point where the curvature is determined and $r_{\rm g} = 2Gm/c^2$ is the gravitational radius (equal to 3 km for the Sun). Averaging (5) over a cell with radius δ yields

$$k = \frac{3}{\delta^3}\int\limits_{r_0}^{\delta} \frac{r_{\rm g}}{r} dr \sim 10^{-46}\ {\rm cm}^{-2}$$

and

$$K_{\rm eff} = -\frac{1}{3}k^2\delta^2 \sim -5 \times 10^{-55}\ {\rm cm}^{-2}.$$

The corresponding radius of curvature is

$$a_{\rm eff} = \sqrt{\frac{1}{|K_{\rm eff}|}} \sim 10^{27}\ {\rm cm}.$$

This quantity exceeds the size of the Galaxy, $L \sim 6 \times 10^{22}$ cm, by at least four orders of magnitude and is of the order of the "radius of the Universe," $l_h \sim 10^{28}$ cm, implying that the contribution of individual stars can be neglected.

2.2. Galaxies

Consider now a Universe filled with galaxies of mass $m_{\rm gal} \sim 10^{11}\, M_\odot$ (the average mass of a galaxy without including dwarf and supergiant galaxies), radius $r_{\rm gal} \sim 1.5 \times 10^{22}$ cm (the average radius of a galaxy), and $\delta_{\rm gal} \sim 5 \times 10^{24}$ cm. The average curvature fluctuation produced by a single galaxy is then

$$k_{\rm gal} = \frac{3}{\delta_{\rm gal}^3}\int\limits_{r_{\rm gal}}^{\delta_{\rm gal}} \frac{r_{\rm g}}{r} dr \sim 4 \times 10^{-57}\ {\rm cm}^{-2}.$$

We now substitute this estimate into (3) for $K_{\rm eff}$ to obtain

$$K_{\rm eff} = -\frac{1}{3}k_{\rm gal}^2\delta_{\rm gal}^2 \sim -10^{-64}\ {\rm cm}^{-2},$$

and the corresponding effective radius of curvature is

$$a_{\rm eff} \sim 10^{32}\ {\rm cm},$$

which exceeds the current radius of the Universe l_h by four orders of magnitude.

2.3. Clusters of Galaxies

Consider now a Universe that is uniformly filled with clusters of galaxies. Since the scatter of their parameters is very large (the mean radius of a cluster is $r_{\rm cl} \sim 1.5 \times 10^{24}$ cm, while the masses of these objects can vary from 10^{13} to $10^{15}\, M_\odot$), we considered two separate cases corresponding to different mean parameters of the galaxy clusters.

(a) We will first suppose that the Universe is filled with poor clusters with masses of the order of $m_{\rm cl} \sim 10^{13}\, M_\odot$ and with the mean intercluster distance being $\delta_{\rm cl} \sim 4 \times 10^{25}$ cm. In this case, the mean curvature fluctuation produced by one such cluster is

$$k_{\rm cl} = \frac{3}{\delta_{\rm cl}^3}\int\limits_{r_{\rm cl}}^{\delta_{\rm cl}} \frac{r_{\rm g}}{r} dr \sim 5 \times 10^{-58}\ {\rm cm}^{-2}.$$

The corresponding effective curvature is then

$$K_{\rm eff} = -\frac{1}{3}k_{\rm cl}^2\delta_{\rm cl}^2 \sim -10^{-64}\ {\rm cm}^{-2},$$

and the effective radius of curvature is

$$a_{\rm eff} \sim 10^{32}\ {\rm cm},$$

which exceeds the radius of the Universe by four orders of magnitude.

(b) A similar model with a Universe populated by rich clusters of galaxies with $m_{\rm cl} \sim 10^{15}\, M_\odot$ with a mean distance between clusters of $\delta_{\rm cl} \sim 10^{26}$ cm yields the curvature fluctuation

$$k_{\rm cl} = \frac{3}{\delta_{\rm cl}^3}\int\limits_{r_{\rm cl}}^{\delta_{\rm cl}} \frac{r_{\rm g}}{r} dr \sim 4 \times 10^{-57}\ {\rm cm}^{-2}.$$

The corresponding effective curvature is

$$K_{\rm eff} = -\frac{1}{3}k_{\rm cl}^2\delta_{\rm cl}^2 \sim -4 \times 10^{-62}\ {\rm cm}^{-2},$$

and the effective radius of curvature is

$$a_{\rm eff} \sim 10^{31}\ {\rm cm},$$

which exceeds the radius of the Universe by three orders of magnitude.

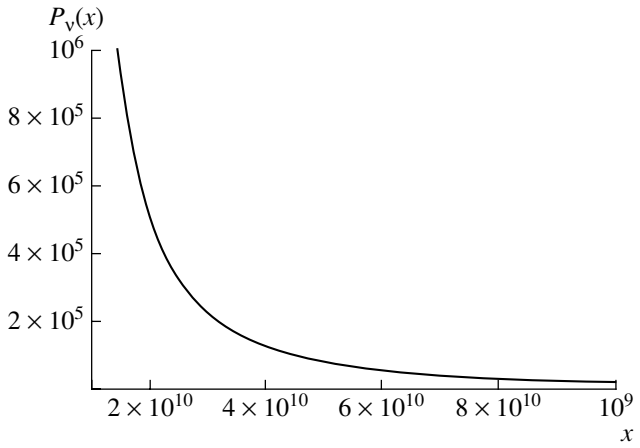

Fig. 1. Density of the distribution of deflection angles $P_\nu(x)$ as a function of the impact parameter p.

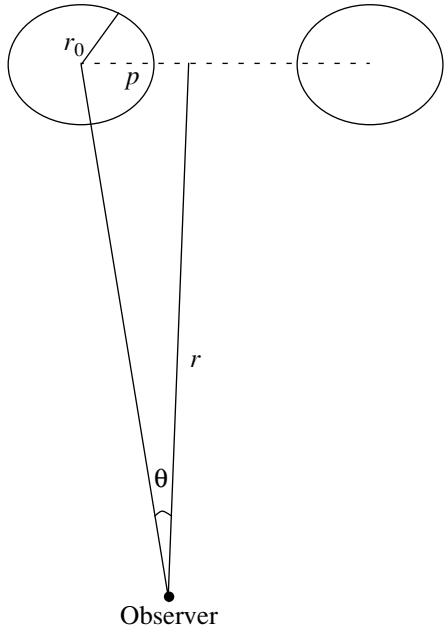

Fig. 2. Impact parameter p for a light beam passing near a star considered to be an inhomogeneity in the Galaxy.

3. DISTANCE BETWEEN CONJUGATE POINTS

3.1. Stars

We are interested in estimating the distance between the conjugate points for an object in whose light cone the inhomogeneities are located. As before, we begin by taking the inhomogeneities to be stars in the Galaxy. We use the following formula [10] to calculate the deflection angle ν of light passing near an inhomogeneity, which appears in (4):

$$\nu = 2r_\mathrm{g}/p. \tag{6}$$

Here, $p = r\theta$ is the impact parameter for light passing between two inhomogeneities that are uniformly distributed over the interval $[R_\mathrm{e}, \delta]$, and $R_\mathrm{e} = \sqrt{2r_\mathrm{g}r}$ is the physical Einstein cone, where r is the distance between the star and the observer. To estimate the distance between the conjugate points, we must know the mean deflection angle $\bar\nu$ averaged over the impact parameters. We can find $\bar\nu$ via the following standard calculations.

Let us assume that the impact parameter p is uniformly distributed over the interval $[R_\mathrm{e}, \delta]$. The distribution of the density $P_\nu(x)$ of the deflection angle (Fig. 1) as a function of the impact parameter p then has the form

$$\begin{cases} 0, & x < \dfrac{2r_\mathrm{g}}{\delta}, \\ P_\nu(x) = \dfrac{2r_\mathrm{g}}{x^2(\delta - R_\mathrm{e})}, & \dfrac{2r_\mathrm{g}}{\delta} \le x \le \dfrac{2r_\mathrm{g}}{R_\mathrm{e}}, \\ 0, & x > \dfrac{2r_\mathrm{g}}{R_\mathrm{e}}. \end{cases} \tag{7}$$

Since the distance r between the star and the observer that appears in the formula for R_e must be much greater than the distance between neighboring stars (Fig. 2), given the size of the Galaxy, it makes sense to set the minimum and maximum values of this distance to the thickness and radius of the optical disk of the Galaxy, respectively, i.e., 1 and 15 kpc, so that the corresponding R_e values are of the order of 4×10^{13} and 2×10^{14} cm, respectively. We now substitute the probability density (7) into the formula for the mean deflection angle ν to obtain

$$\bar\nu = \int\limits_{\frac{2r_\mathrm{g}}{\delta}}^{\frac{2r_\mathrm{g}}{R_\mathrm{e}}} P_\nu(x)x\,dx = \frac{2r_\mathrm{g}}{\delta - R_\mathrm{e}} \ln\frac{R_\mathrm{e}}{\delta} \sim 10^{-7\prime\prime}. \tag{8}$$

The standard deviation is equal to

$$\Delta\nu = \left(\int\limits_{\frac{2r_\mathrm{g}}{\delta}}^{\frac{2r_\mathrm{g}}{R_\mathrm{e}}} P_\nu(x)(x-\bar\nu)^2 dx \right)^{1/2} \tag{9}$$

$$= \left(\frac{4r_\mathrm{g}^2}{\delta R_\mathrm{e}} + \frac{16r_\mathrm{g}^2 \ln^2\dfrac{R_\mathrm{e}}{\delta}}{(\delta - R_\mathrm{e})^2} \right)^{1/2} \sim 10^{-6\prime\prime}.$$

We then substitute the values of r_g, δ, and R_e into (8) and (9) to obtain $\bar\nu \sim 10^{-7}$ arcsec and $\Delta\nu \sim 7\times10^{-6}$ arcsec for $r = 1$ kpc and $\bar\nu \sim 10^{-7}$ arcsec and $\Delta\nu \sim 3\times4\times10^{-6}$ arcsec for $r = 15$ kpc.

Substituting the above $\bar\nu$ value into (4) yields

$$\Delta l \sim 10^{32}\ \mathrm{cm}.$$

A comparison of the distance Δl between the conjugate points and the radius of the Universe L shows that

$$\Delta l/l_h \sim 10^3,$$

implying that the conjugate points are clearly not observable at the present stage of the evolution of the Universe.

3.2. Galaxies

We can perform similar calculations for the case when galaxies are taken to be the density inhomogeneities. A light beam passing near a galaxy is deflected by the mean angle $\bar{\nu}$ equal to

$$\bar{\nu} = \frac{2r_\mathrm{g}}{\delta_\mathrm{gal} - R_\mathrm{e}} \ln \frac{R_\mathrm{e}}{\delta_\mathrm{gal}} \sim 0.02'',$$

and the standard deviation of this angle is

$$D(\nu) = \left(\frac{4r_\mathrm{g}^2}{\delta_\mathrm{gal} R_\mathrm{e}} + \frac{16 r_\mathrm{g}^2 \ln^2 \dfrac{R_\mathrm{e}}{\delta_\mathrm{gal}}}{(\delta_\mathrm{gal} - R_\mathrm{e})^2} \right)^{1/2} \sim 0.08''.$$

We now substitute the mean deflection angle $\bar{\nu}$ into (4) and find the distance between the conjugate points for an object with galaxies in its light cone to be

$$\Delta l \sim 10^{32} \text{ cm},$$

which, as in the case of stars, exceeds the radius of the Universe l_h by four orders of magnitude and is therefore not currently observable.

3.3. Clusters of Galaxies

To estimate the distance between the conjugate points for the case when the inhomogeneities are taken to have the form of clusters of galaxies, we must substitute the corresponding values for clusters of galaxies into (8) and (9).

(a) If we assume that the Universe is uniformly filled with poor clusters of galaxies with masses of the order of $m_\mathrm{cl} \sim 10^{13}\, M_\odot$ separated by a mean intercluster distance of $\delta_\mathrm{cl} \sim 4 \times 10^{25}$ cm, the mean deflection angle $\bar{\nu}$ for light passing in the vicinity of a cluster of galaxies is of the order of

$$\bar{\nu} = \frac{2r_\mathrm{g}}{\delta_\mathrm{gal} - R_\mathrm{e}} \ln \frac{R_\mathrm{e}}{\delta_\mathrm{gal}} \sim 0.2'',$$

with its standard deviation being

$$D(\nu) = \left(\frac{4r_\mathrm{g}^2}{\delta_\mathrm{gal} R_\mathrm{e}} + \frac{16 r_\mathrm{g}^2 \ln^2 \dfrac{R_\mathrm{e}}{\delta_\mathrm{gal}}}{(\delta_\mathrm{gal} - R_\mathrm{e})^2} \right)^{1/2} \sim 0.6''.$$

Substituting the mean deflection angle $\bar{\nu}$ into (4) indicates that the mean distance between the conjugate points is

$$\Delta l \sim 10^{32} \text{ cm},$$

which also exceeds the radius of the Universe l_h by four orders of magnitude and is therefore not observable at the present epoch.

(b) In a similar way, for the case of a Universe uniformly filled with rich clusters of galaxies with $m_\mathrm{cl} \sim 10^{15}\, M_\odot$ separated by a mean intercluster distance of $\delta_\mathrm{cl} \sim 10^{26}$ cm, we find that the mean deflection angle $\bar{\nu}$ for a light beam passing in the vicinity of such a cluster of galaxies is

$$\bar{\nu} = \frac{2r_\mathrm{g}}{\delta_\mathrm{gal} - R_\mathrm{e}} \ln \frac{R_\mathrm{e}}{\delta_\mathrm{gal}} \sim 4'',$$

with its standard deviation being

$$D(\nu) = \left(\frac{4r_\mathrm{g}^2}{\delta_\mathrm{gal} R_\mathrm{e}} + \frac{16 r_\mathrm{g}^2 \ln^2 \dfrac{R_\mathrm{e}}{\delta_\mathrm{gal}}}{(\delta_\mathrm{gal} - R_\mathrm{e})^2} \right)^{1/2} \sim 9''.$$

Substituting the mean deflection angle $\bar{\nu}$ into (4), we find that the mean distance between the conjugate points for an object with clusters of galaxies in its light cone is

$$\Delta l \sim 10^{31} \text{ cm},$$

which also exceeds the radius of the Universe l_h.

4. DISCUSSION

Our estimates for the effective radius of curvature and the mean distance between the conjugate points indicate that neither of the effects considered can be observed at the present stage of evolution of the Universe. However, we consider it useful to identify the parameters of the inhomogeneities that we have considered here for which the Zel'dovich effect would become observable.

Let us first consider the estimates of the effective radius of curvature. To determine the ratios of the inhomogeneity parameters m and δ for which the corresponding effective radius does not exceed the critical value (equal to the size of the Galaxy for stars, $L \sim 6 \times 10^{22}$ cm, and the radius of the Universe, $l_h \sim 10^{28}$ cm, for galaxies and clusters of galaxies), we set a_eff equal to the critical value and then derived functions relating m and δ; these allow us to determine pairs of m and δ values that correspond to the possibility of observing the Zel'dovich effect. If the inhomogeneities are taken to be stars in the Galaxy, setting the critical value of a_eff equal to the

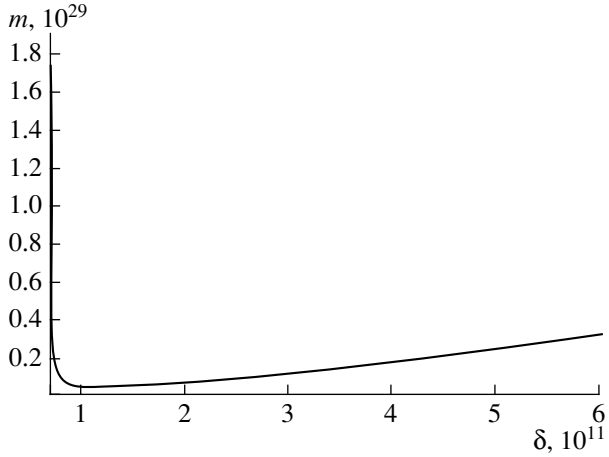

Fig. 3. Form of the function $m(\delta)$ for the case when the inhomogeneities are taken to be stars in the Galaxy.

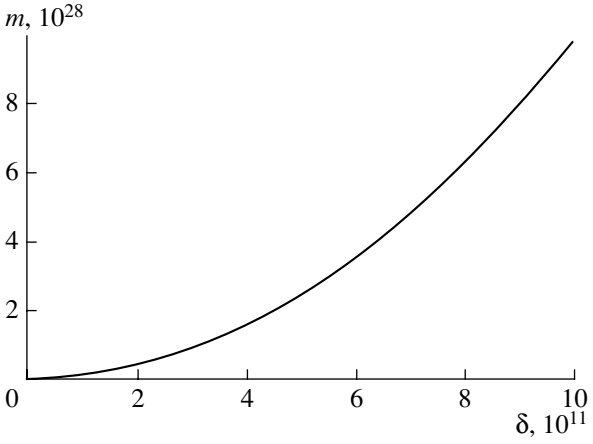

Fig. 4. Form of the function $m(\delta)$ for the case of a Universe uniformly filled with galaxies.

size of the Galaxy, $L \sim 6 \times 10^{22}$ cm, yields the following function $m(\delta)$, which includes the size of the inhomogeneities r_0 as a parameter:

$$m(\delta) = 2 \times 10^5 \text{ g/cm}^2 \frac{\delta^2}{\ln(\delta/r_0)}.$$

(Here and below, m is in grams, δ is in centimeters, and the coefficient of δ^2 is dimensioned.) As is evident from the above, this function has a singularity at $\delta = r_0$. We set r_0 in this formula equal to the radius of the star to derive a specific form of the function $m(\delta)$ (Fig. 3), which can be used to identify pairs of m and δ values that satisfy the condition $a_{\text{eff}} \leq L$. It is obvious that this condition is satisfied by all pairs of m and δ that lie on or below the plot. Taking the mass of a star to be $m_{\text{star}} = M_\odot$, the corresponding distance on the plot is $\delta \sim 3 \times 10^{15}$ cm, while the mean distance between the stars in the Galaxy is of the order of 10^{19} cm, which is four orders of magnitude greater. Thus, in order for the contribution of individual stars to the Zel'dovich effect to be observable, the distance between neighboring stars would have to be a factor of about 10^4 smaller than the current value. Alternatively, with the current distances between stars, their masses would have to be of the order of $10^{10}\, M_\odot$.

For a Universe uniformly filled with galaxies, we set the critical a_{eff} equal to $l_h \sim 10^{28}$ cm. Setting r_0 equal to the mean size of galaxies yields the function

$$m(\delta) = 0.2 \text{ g/cm}^2 \frac{\delta^2}{\ln(\delta/1.5 \times 10^{22})},$$

whose form is identical to that shown in Fig. 4. In order for the Zel'dovich effect to be observable in the case of a Universe that is uniformly filled with galaxies, the mean distance between galaxies (with $m_{\text{gal}} \sim 10^{11}\, M_\odot$) would have to be about a factor of 100 smaller than the current distance between neighboring galaxies; alternatively, with $\delta_{\text{gal}} \sim 5 \times 10^{24}$ cm, the masses of galaxies would have to be of the order of $10^{15}\, M_\odot$.

Analogous calculations for a Universe that is uniformly filled with clusters of galaxies yield the formula

$$m(\delta) = 0.2 \text{ g/cm}^2 \frac{\delta^2}{\ln(\delta/1.5 \times 10^{24})},$$

which determines pairs of m and δ values that correspond to the possibility of observing the Zel'dovich effect for clusters of galaxies. This is possible, in particular, with the pairs $m \sim 10^{13}\, M_\odot$ and $\delta \sim 10^{23}$ cm or $m \sim 10^{17}\, M_\odot$ and $\delta \sim 10^{25}$ cm.

Let us now consider the estimates of the distance between the conjugate points, Δl. In the case of a Galaxy that is uniformly filled with stars, setting the critical $\Delta l = L$ yields

$$m(\delta) = 10^5 \text{ g/cm}^2\, \delta^2,$$

which determines pairs of m and δ values that correspond to the possibility of observing the Zel'dovich effect (Fig. 4). Examples of pairs of m and δ values are $m \sim M_\odot$ and $\delta \sim 10^{14}$ cm, or $m \sim 10^{10}\, M_\odot$ and $\delta \sim 10^{19}$ cm.

For the case of a Universe uniformly filled with galaxies or clusters of galaxies, the critical value is equal to $\Delta l = l_h$, and the corresponding function $m(\delta)$ has the form

$$m(\delta) = 1 \text{ g/cm}^2\, \delta^2.$$

The Zel'dovich effect can be observed in the case of the pairs of m and δ values $m \sim 10^{11}\, M_\odot$ and $\delta \sim 10^{22}$ cm, or $m \sim 10^{15}\, M_\odot$ and $\delta \sim 10^{24}$ cm.

5. CONCLUSIONS

In this paper, we have analyzed two effects due to small-scale deviations of the Universe from uniformity and isotropy. The first effect consists of a decrease in the curvature of the spatial cross section and makes a Universe with its mean density equal to the critical density appear to an observer like a world whose spatial cross section has negative curvature. We also analyzed the cumulative effect of many randomly located density perturbations, which results in the appearance of conjugate points. Our computations neglected the details of the structure or any specific features of inhomogeneities (e.g., clustering of galaxies), since the resulting estimates show that the magnitude of the effects considered is extremely small, and including more detailed information about the properties of the inhomogeneities would not appreciably change the results. We estimated the effective curvature and radius of curvature, as well as the mean distance between the conjugate points when the density inhomogeneities are taken to be stars in our Galaxy, galaxies, and clusters of galaxies. The estimates obtained show that the magnitude of the effects considered is extremely small, and that these effects are not observable at the present stage of evolution of the Universe.

Note, however, that, despite its small magnitude, the Zel'dovich effect maintains its importance for cosmology. The reason is that the difference between spaces with positive, negative, and zero curvature is not only quantitative but also qualitative in nature. If the mean density ρ of the Universe indeed coincides with the critical density ρ_{cr}, $\bar{K} = 0$, even a small negative addition could change the properties of the Einstein equations radically.

Zel'dovich suggested that the Einstein equations be solved according to the following scheme. Solutions are first constructed for the Einstein equations for the mean curvature. These solutions are then supplemented with the curvature perturbations corresponding to individual objects. The results obtained suggest that it should be possible to somehow take the perturbations into account in the Einstein equations and solve the perturbed equations. However, we cannot suggest any way to correctly include the perturbations in the equations and subsequently solve them. These operations are believed to have no effect on the rate of expansion of the Universe. The qualitative difference between flat and Lobachevski geometries may indicate that this picture requires some correction.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 04-02-16094 and 04-02-17288). We are also very grateful to M.V. Sazhin and D.D. Sokolov for discussions and useful comments made during the preparation of this paper.

REFERENCES

1. Ya. B. Zel'dovich, Astron. Zh. **41**, 19 (1964) [Sov. Astron. **8**, 13 (1964)].
2. Ya. B. Zel'dovich and I. D. Novikov, *Relativistic Astrophysics* (Nauka, Moscow, 1967; Univ. of Chicago Press, Chicago, 1971).
3. V. G. Lamburt, D. D. Sokolov, and V. N. Tutubalin, Mat. Zametki **74**, 416 (2003).
4. M. V. Sazhin, V. E. Zharov, A. V. Volynkin, and T. A. Kalinina, Mon. Not. R. Astron. Soc. **300**, 287 (1998).
5. V. E. Zakharov and M. V. Sazhin, Usp. Fiz. Nauk **168**, 1041 (1998) [Phys. Usp. **41**, 945 (1998)].
6. V. E. Zharov, M. V. Sazhin, and N. A. Chuĭkova, Astron. Zh. **77**, 144 (2000) [Astron. Rep. **44**, 122 (2000)].
7. M. Sazhin, Gravit. Cosmol. **6**, 214 (2000).
8. M. V. Sazhin, V. E. Zharov, and T. A. Kalinina, Mon. Not. R. Astron. Soc. **323**, 952 (2001).
9. V. G. Lamburt, Mat. Zametki (2003) (in press).
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics,* Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1988; Pergamon, Oxford, 1975).

Translated by A. Dambis

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 777–782. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 874–880.

A Strong Flare of the H_2O Maser in Sagittarius B2(M)

O. Ramírez Hernández[1], E. E. Lekht[2], and A. M. Tolmachev[3]

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

[2] *Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis Enrique Erro No. 1, Apdo Postal 51 y 216, 72840 Tonantzintla, Puebla, Mexico*

[3] *Pushchino Radio Astronomy Observatory, Astro Space Center, Lebedev Institute of Physics, Pushchino, Moscow oblast, 142290 Russia*

Received February 7, 2005; in final form, May 18, 2005

Abstract—Results of a study of a strong flare of H_2O maser emission in the star-forming region Sgr B2(M) in 2004 are reported. The observations were carried out on the 22-m radio telescope of the Pushchino Radio Astronomy Observatory. The main emission, with its flux density reaching 3800 Jy, was concentrated in a narrow radial-velocity interval (about 3 km/s) and was most likely associated with the compact group r, while the emission at $V_{LSR} > 64$ km/s came from group q. After 1994, the variations of the H_2O maser emission in Sgr B2(M) became cyclic with a mean period of 3 years. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Sagittarius B2 is one of the most complex and active star-forming regions in our Galaxy. It includes more than ten separate HII regions, located in three main star-forming centers: Sgr B2(N), Sgr B2(M), and Sgr B2(S) [1]. The structure of each is rather complicated. The largest centers of star formation are N and M, with region N believed to be the younger one [2]. H_2O (and other) maser emission is observed toward both regions. Lekht *et al.* [3] have shown that the total emission of each source alternatively dominates over that of the other. The emission from the main source (M) was dominant at the end of 2003.

The main H_2O maser emission toward Sgr B2(M) is associated with a compact group of maser spots. The spots are seen against an ultracompact HII F region and, according to Kobayashi *et al.* [4], cover an area of 0.04 pc × 0.04 pc (for a distance to Sgr B2 of 8.5 kpc). It is most likely that the HII region F is the activity center for the OH, H_2CO, and H_2O masers. Gaume *et al.* [5] have shown that this region contains of four subsources, F1–F4. The diameter of each is about 0.01 pc, and the brightness temperature is no lower than 23 400 K [6], whereas the normal temperatures of HII regions in the Galaxy are 6000–9000 K. The radial velocity derived from the H76α radio recombination line from the entire source F is ≈60 km/s [7].

The strongest H_2O emission from Sgr B2(M) appeared at velocities of 60–70 km/s. This group of maser spots is located between subsources F1–F4 [8]. A molecular outflow in the Sgr B2(M) region [8, 9] is also associated with ultracompact HII regions and maser sources [3]. It cannot be excluded that the water maser is located within the molecular outflow, not at the edge of some of the F regions. The most intense H_2O emission of Sgr B2(M) is most likely associated with this outflow.

2. OBSERVATIONS AND DATA ANALYSIS

Our observations of the H_2O maser source toward Sgr B2 ($\alpha_{1950} = 17^h44^m10^s$, $\delta_{1950} = -28°22'00''$) were carried out on the 22-m radio telescope of the Pushchino Radio Astronomy Observatory in 2004. In the observations at low elevations, the system noise temperature was 150–250 K. The antenna beamwidth at 1.35 cm is 2.6′. For a pointlike unpolarized source, an antenna temperature of 1 K corresponds to a flux density of 25 Jy. The signal was processed by a 128-channel filter-bank spectrum analyzer with a resolution of 7.5 kHz (0.101 km/s at the radial velocity of the 1.35-cm line).

Figure 1 presents the H_2O spectra toward Sgr B2 (component M), corrected for the absorption in the Earth's atmosphere. The horizontal axis plots the radial velocity with respect to the Local Standard of Rest, and the vertical axis plots the flux density in Janskys. The vertical arrow shows the scale in Janskys, and the horizontal lines show the spectrum baselines.

Sgr B2 hosts two main groups of maser spots, associated with the components N and M and separated in declination by 45″ [4]. Since the beamwidth of the 22-m (2.6′) telescope is larger than this, we

1063-7729/05/4910-0777$26.00

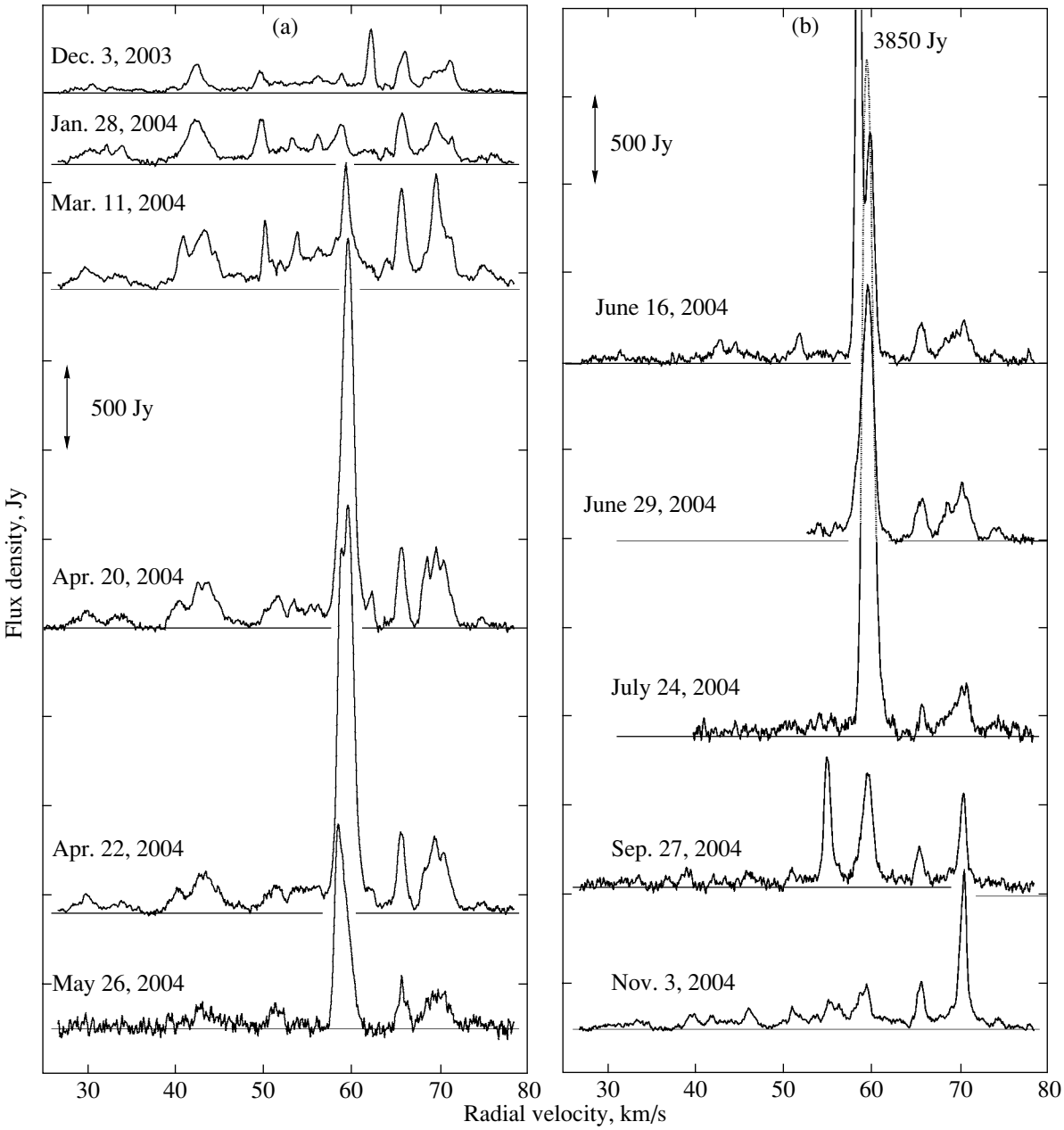

Fig. 1. Spectra of the H_2O maser emission of Sgr B2 obtained in 2004. The vertical arrows show the scale in Janskys. The horizontal axis plots the radial velocity (km/s) relative to the Local Standard of Rest.

made observations at four points along the line connecting the sources M and N in order to determine which emission features belonged to which source. Figure 2 presents the results of these observations at different stages of the flare evolution. The intervals between the points are 45″. The fluxes at different radial velocities are shown as functions of the antenna pointing declination. All the curves are numbered. In addition, radial velocities are given for most of them. For convenience, the fluxes of features 3 and 4 in Fig. 1b have been reduced by a factor of three, and those of feature 4 in Fig. 1c by a factor of two. Scans with maxima close to source M are plotted as curves with closed circles, and those with maxima close to source N with open circles. At radial velocities where the fluxes from M and N were more or less equal, the scan maxima are at an intermediate location; these are labeled with asterisks.

A strong flare of the emission took place in the radial-velocity interval 57.5–62.5 km/s. This part of the spectrum is shown in Fig. 3, which plots all the spectra for 2004, numbered in chronological order.

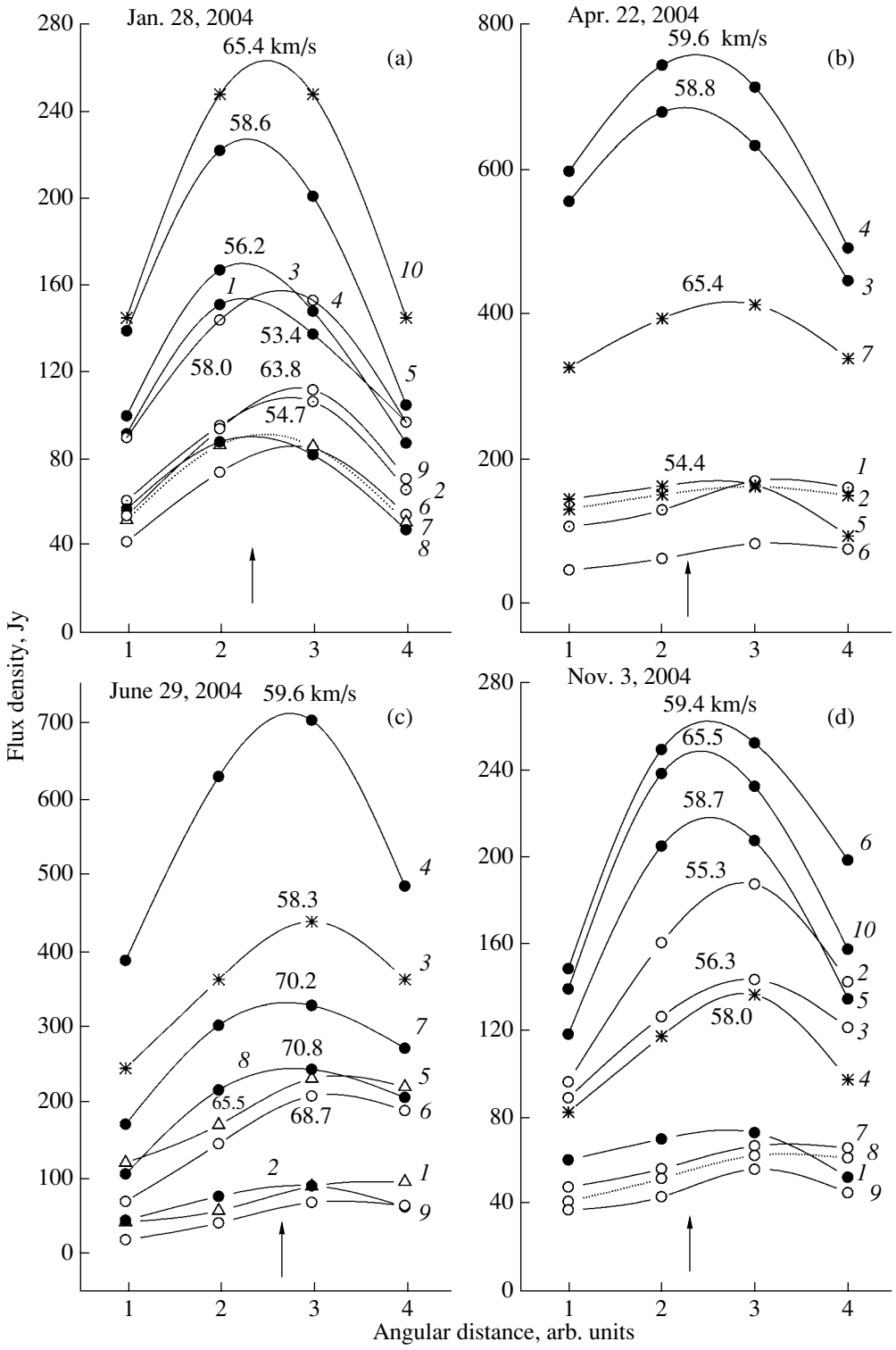

Fig. 2. The fluxes as a function of the antenna pointing declination. One horizontal scale unit corresponds to 45″. The vertical arrow shows the position of Sgr B2(M).

The correspondence of the spectra to the epochs of the observations is given at the right. In this radial-velocity interval, we have separated the emission into individual spectral components, whose flux variations are shown in Fig. 4. As we can see from Fig. 2, the most intense emission features are identified with the main source, Sgr B2(M). This suggests that the strong flare happened precisely in this source.

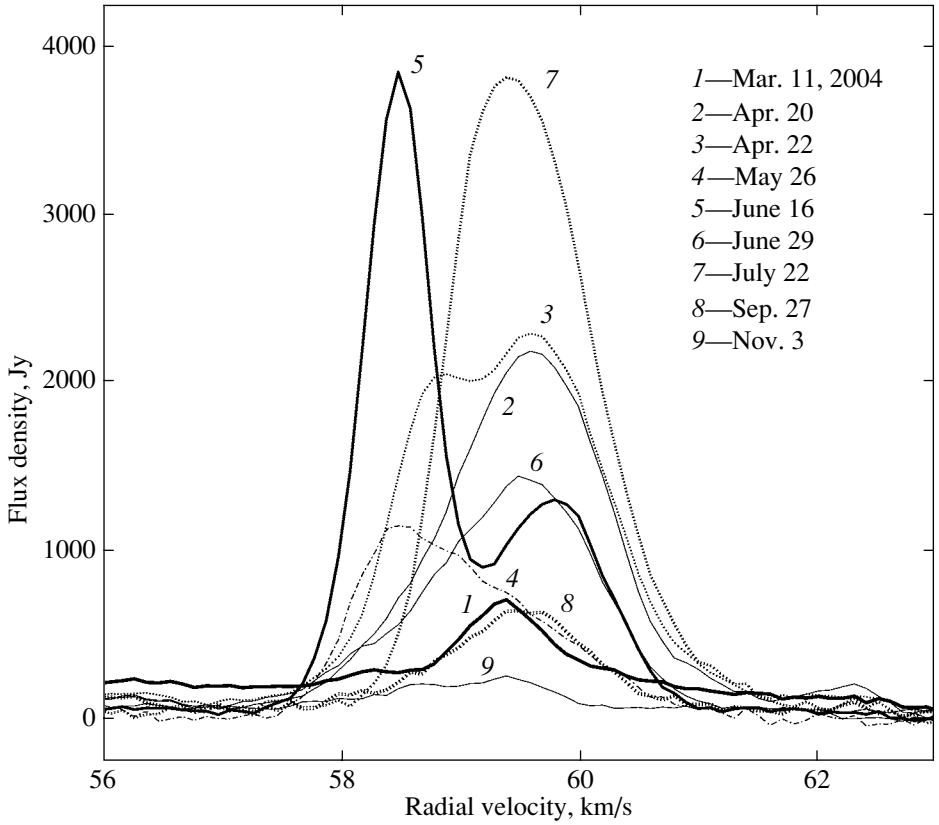

Fig. 3. Overlay of the central parts of the H_2O spectra obtained in 2004. Epochs are labeled at the right.

We observed fast variability in the emission feature at 58.8 km/s, whose flux increased from 800 to 2000 Jy in two days, with a mean growth rate of about 25 Jy/h. Small variations of the spectrum's structure occurred during the observations of April 22 aimed at identifying the sources of the spectral features (the four-point observations), which lasted almost 1.5 h. Since the feature at 59.6 km/s varied insignificantly, we adopted it as a reference. We normalized the four consecutive spectra measured at the four points to this reference feature. Figure 5 shows the central parts of the normalized spectra.

The variations of the emission at 58.8 km/s are within 4%, which corresponds to 80 Jy on an absolute flux scale. The flux-variation rate was about 50 Jy/h. This coincidence is rather good, in view of the possibly unsteady rate of the flux variation. In addition, the flux of the reference feature could have changed slightly. Depending on the phase of these changes, the rate of the flux variation of the feature at 58.8 km/s could either increase or decrease.

The fast flux variability we have detected is real, since it is based on relative measurements. We have eliminated all errors due to calibration and variations in the weather conditions during the observations, which can strongly affect the absolute flux magnitude, especially at low source elevations. The difference in the curves near velocities of 57.5 and 62 km/s is due to either noise fluctuations or the presence of emission from the northern source, as can be seen in Fig. 2b.

We have also plotted the flux variations of other emission features outside the central part of the H_2O spectrum (Fig. 6). This is necessary to search for possible correlations of the flare with variations in other spectral features and to estimate the scale of the maser flare (local or global).

3. DISCUSSION

3.1. Cyclic Activity of the Maser

Lekht *et al.* [3] showed that the emission of either the northern (N) or the main (M) source alternatively dominates. Since the end of 2003, the source M has become the most intense. The strong flare in 2004, which lasted almost eight months, is associated with precisely this source.

The alternating domination of the emission from each of the two sources could be a consequence of activity cycles in at least one of the two H_2O masers. Using the results of the monitoring for 1994–2004, we can estimate the period of such activity only for

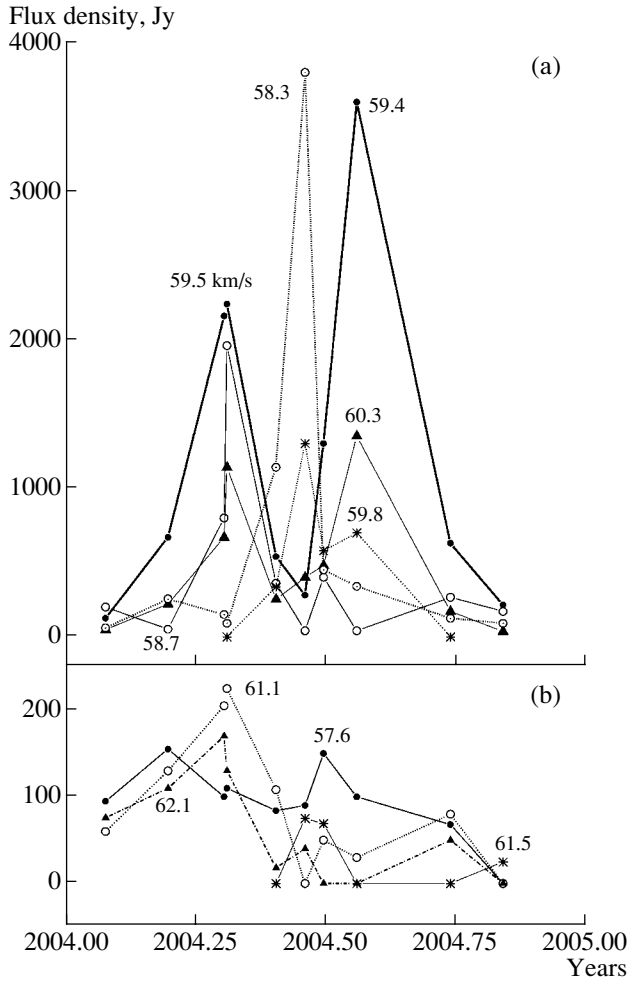

Fig. 4. Flux and radial-velocity variations of the emission features at 57.5–62.5 km/s.

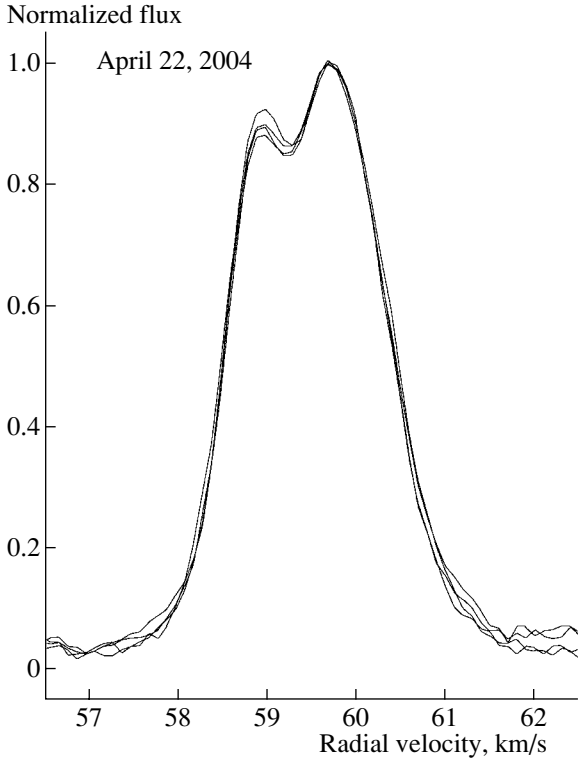

Fig. 5. Rapid flux variability of the emission feature at 58.8 km/s with respect to the emission at 59.6 km/s (see text).

the main source (M), which is about 3 years. We distinguished only three activity cycles for the northern source separated by two and three years in this same interval.

In addition, we observed shorter lived flares of activity lasting several months in both sources. Such flares happened throughout our monitoring of 1982–2004, with the mean interval between them being less than 0.8 yr.

3.2. Identification of the Emission Features

The flare of 2004 in the source M was only slightly weaker than the strongest flare that happened in Sgr B2 in 1986, which was associated with the northern source (N). There are important differences between these flares. During the 1986 flare, the maser emission was consecutively excited in order of increasing radial velocities in the interval 47–75 km/s. The 1986 flare was identified with an outflow of material from a rotating accretion disk in the source Sgr B2(N) [10].

In the 2004 flare in the source M, the most intense rapidly varying emission appeared in a narrow velocity interval that approximately matches the thermal linewidth. The emission maxima of individual features virtually coincided.

In 2004, the mean velocity of the strongest emission coincided with the central velocity of the H76α radio recombination line, 60.3 km/s [7], in the vicinity of the HII region F. The position of this group of maser spots coincides with the dense part of the 60-km/s compact molecular cloud. According to Kobayashi *et al.* [4], the intense maser emission at radial velocities of 55–70 km/s forms two groups of maser spots, r and q. We can assume that the strong flare at 58–61 km/s was associated with group r.

The different characters of the 1986 and 2004 flares may be due to differences in the structures of the sources M and N. The common feature is that, as in the northern source, the maser spot clusters r and q in M are associated with a molecular outflow.

3.3. Correlation and Anticorrelation of the Emission

The 2004 flare can be represented as two subflares. During the first, the flux densities from the

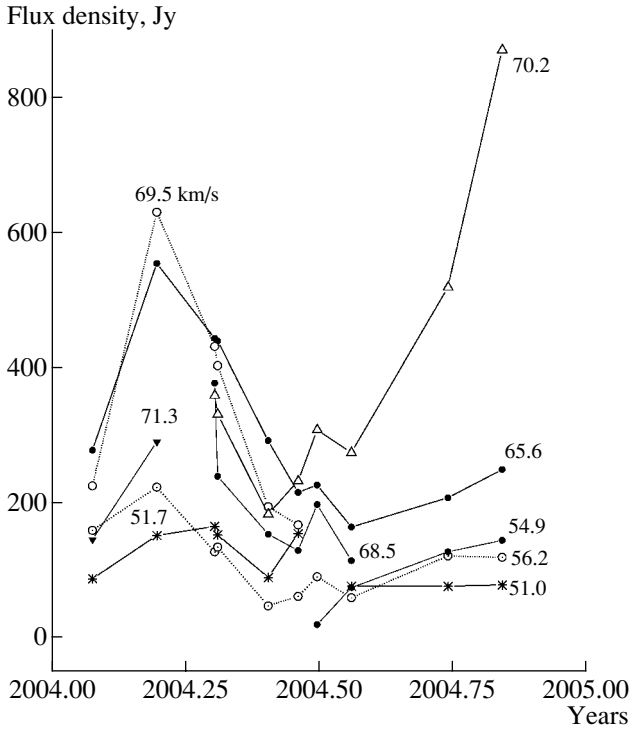

Fig. 6. Flux variations of various emission features.

two main features, at 59.5 and 60.3 km/s, exceeded 2300 and 1200 Jy, while during the second, they exceeded 3500 and 1400 Jy (Fig. 4a). The subflares were separated by some time interval, in which strong emission at 58.3 km/s (3800 Jy) and weaker emission at 59.8 km/s were observed. The emission in other features was strongly suppressed at that time. Obviously, the fluxes of the features with close radial velocities were anticorrelated. A similar evolution took place in three other features, at 61.1, 61.5, and 62.1 km/s. The emission in the lateral features varied in antiphase with the central feature (Fig. 4b).

There is a good coincidence of the emission maxima of individual emission features. Since the flare was fairly short-lived, the time intervals between consecutive observing sessions turned out to be insufficient to accurately determine the epochs of the emission maxima of individual features. The upper limit for the time shift between the maxima does not exceed one month. This result also suggests that the cluster of maser spots that are responsible for the emission at 57.5–61 km/s is fairly compact.

The fluxes of the emission features outside 57.5–61 km/s are also correlated, and their maxima nearly coincide. The mean peak falls in March 2004, anticipating the maximum of the main group of features, which occurs in April. According to Kobayashi *et al.* [4], the emission at $V_{LSR} > 64$ km/s may belong to maser spot q, which lies slightly closer to the center of the compact HII region than the cluster r.

4. CONCLUSION

We have studied a strong flare in Sgr B2 that occurred 18 years after a similar flare in 1986 [3]. The flares took place in different sources (the northern and central components of Sgr B2) and display substantial differences. Let us list the main results we have obtained.

(1) The strong flare of 2004, which took place in the main source M of the Sgr B2 region, is most likely associated with the compact group of maser spots r [4].

(2) The cyclic activity of the water maser in Sgr B2(M) has a mean period of about three years.

(3) The radial velocities of the main features of the flare were within the thermal linewidth, possibly testifying to the compactness of the spot cluster.

(4) The flare consisted of two subflares. This structure of the flare is due to an anticorrelation between the fluxes of two pairs of emission features with similar radial velocities.

ACKNOWLEDGMENTS

Work on the RT-22 Radio Telescope is supported by the Ministry of Science and Education of the Russian Federation (registration no. 01-10). The authors are grateful to the staff of the Pushchino Radio Astronomy Observatory for help with the observations.

REFERENCES

1. J. M. Benson and K. J. Johnston, Astrophys. J. **277**, 181 (1984).
2. S. Y. Liu, D. M. Mehringer, Y. Miao, and E. Snyder, Astrophys. J. **501**, 680 (1998).
3. E. E. Lekht, O. Ramírez-Hernández, and A. M. Tolmachev, Astron. Zh. **81**, 1059 (2004) [Astron. Rep. **48**, 965 (2004)].
4. H. Kobayashi, M. Ishiguro, Y. Chikada, *et al.*, Publ. Astron. Soc. Jpn. **41**, 141 (1989).
5. R. A. Gaume, M. J. Claussen, C. G. De Pree, *et al.*, Astrophys. J. **449**, 663 (1995).
6. C. G. de Pree, R. A. Gaume, W. M. Goss, and M. J. Claussen, Astrophys. J. **464**, 788 (1996).
7. P. R. Roelfsema, W. M. Goss, J. B. Whiteoak, *et al.*, Astron. Astrophys. **175**, 219 (1987).
8. Y. J. Kuan and L. E. Snyder, Astrophys. J. **470**, 981 (1996).
9. D. C. Lis and P. F. Goldsmith, Astrophys. J. **356**, 195 (1990).
10. E. E. Lekht, O. Ramírez-Hernández, A. M. Tolmachev, and I. I. Berulis, Astron. Zh. **81**, 195 (2004) [Astron. Rep. **48**, 171 (2004)].

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 783–800. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 881–899.

Light-Curve Syntheses for Close Binary Systems: Modeling Spiral Waves in an Elliptical Disk around a White Dwarf

T. S. Khruzina

Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia

Received December 18, 2004; in final form, February 17, 2005

Abstract—We present an algorithm for synthesizing the light curve of a close binary consisting of a normal star (a red dwarf that fills its Roche lobe) and a spherical star (a white dwarf). The spherical component is surrounded by an elliptical accretion disk with a complex shape: it is geometrically thin near the spherical star and geometrically thick at the edge of the disk. An additional complication is presented by the presence of a one- or two-armed spiral pattern at the inner surface of the disk. The maximum height of the spiral arm above the disk surface is located at $\sim 0.9\,R_d$, and the height decreases exponentially as the arm approaches the inner regions of the disk. Shielding of the inner hot parts of the disk by the crests of the spirals results in the formation of "steps" in out-of-eclipse parts of the orbital light curves. The algorithm takes into account the presence of a "hot line" by the lateral surface of the disk, making it possible to model binary systems in both quiescence and outburst. In the latter case, the hot line degenerates into a small bulge at the outer lateral surface of the disk, which can be considered an analog of a hot spot. The algorithm was applied to the orbital light curve of the cataclysmic binary IP Peg during its October 30, 2000, outburst. To explain the variations of the out-of-eclipse brightness of the system during the outburst, it is necessary to include the presence of a one-armed spiral wave at the inner surface of the disk, close to the periastron of the elliptical disk. We have obtained the parameters of IP Peg during the outburst for various models of the system.

1. INTRODUCTION

For the last four decades, the hot-spot model suggested by Gorbatskii [1] and Smak [2] in the late 1960s has been used to interpret the light curves of cataclysmic variables. In this model, the interaction between a stream of matter and the accretion disk is accompanied by the formation of a shock wave, and a region of energy release forms where the stream comes into contact with the disk, called a "hot spot." This model could explain the formation of a hump in the light curves of cataclysmic variables close to orbital phases $\phi \sim 0.7$–0.85. However, as observational data accumulated, deviations from this model were revealed. Many observed features of the light curves could not easily be explained, such as anomalous light curves that had fluxes higher in the egress from eclipse than in the ingress and secondary humps in the out-of-eclipse light curve. Attempts to save the hot-spot model by considering elliptical rather than circular disks did not resolve these problems, although they improved the agreement between observed and synthetic light curves (an algorithm for synthesizing light curves for the case of an elliptical disk with a hot spot on its lateral surface is presented in [3]).

Three-dimensional gas-dynamical computations of the mass flows in close binary systems have revealed that, in self-consistent solutions for various initial conditions, independent of the assumed temperature of the gas in the outer parts of the disk, there is no hot spot, i.e., no shock interaction between the stream of matter flowing from the inner Lagrangian point and the accretion disk [4–7]. The resulting region of enhanced energy release is located outside the disk and is due to the interaction of the circumdisk halo and intercomponent envelope with the stream. It was proposed to call this region a "hot line" [4–7]. Comparisons of synthetic light curves obtained for this model and observed light curves of selected cataclysmic variables and binary X-ray sources have shown excellent agreement [8–12]. The hot-line model was able to reproduce many details of the out-of-eclipse parts of the light curves that remained unexplained in the hot-spot model. However, the hot-line model was much less successful for light curves obtained during outburst, for which there remained details that could not be reproduced. In particular, among eclipsing close binaries, the light curves obtained during the outbursts of cataclysmic variables and superoutbursts of SU UMa-type systems contain dips at phases ~ 0.25 and ~ 0.75, which may be related to an increase in the geometrical thickness of the disk at these phases. Similar dips at phases ~ 0.2

1063-7729/05/4910-0783$26.00

and ~0.8 are also observed in the ultraviolet. The hot-line model is not able to explain such dips.

In recent years, observational data have provided evidence for the existence in some close binaries (such as IP Peg, U Gem, EX Dra, SS Cyg, V347 Pup) of two-armed spiral shocks in the outer parts of their accretion disks during outbursts [13]. Although Doppler tomography provides some evidence for a certain amount of asymmetry in the accretion disks in quiescence, it definitely rules out the presence of such spiral arms. Spiral pattern features are also absent from systems with high component-mass ratios. For example, Doppler tomography of OY Car in outburst [14] shows the presence of an extended spiral arm along the side of a gaseous stream that is undergoing a shock interaction, but does not show two spiral shocks.

High-phase-resolution spectroscopy of cataclysmic variables during outbursts that enables mapping of the disks has been carried out for only a small number of objects. However, the available data provide evidence that these phenomena have a common physical nature: the most probable origin of such large-scale asymmetries is tidal shocks initiated by the presence of a fairly massive secondary component in the system.

Analysis of the main processes heating and cooling the matter in the accretion disks of close binaries [15, 16] shows that the gas temperature in the outer parts of the disk is from ~10 000 K ("cold gas") to ~1 000 000 K ("hot gas"). Three-dimensional calculations with hot disk gas (see, e.g., [17]) show that a one-armed spiral shock forms close to the apoastron of the accretion disk; the flow structure in the vicinity of the disk periastron, where the second arm should be located, is determined by the stream of matter from the inner Lagrangian point, which, apparently, prevents the formation of the second spiral arm. Calculations of cool gas flows in close binaries [18] show that, in this case, the disk is more circular and denser, and the height of its outer edge is lower than in the case of a disk with hot gas; in such a disk, a second arm of the tidal spiral forms near the disk periastron. Both arms are located in the outer regions of the disk, since the stream of matter does not strongly influence the inner dense parts of the disk, and the tidal spiral arms do not propagate into its inner regions.

Mapping of the eclipses of SU UMa stars [19] also shows the presence of three sources of emission in the outer regions of the disk. The locations of these sources coincide with the two arms of a tidal spiral shock and a shock identified with the hot line. The linear scales of these features are relatively large. As we noted above, observations of outbursts and superoutbursts of eclipsing cataclysmic variables show dips in the light curves at phases ~0.2–0.25 and ~0.75–0.8, in both the optical and ultraviolet. In the model considered, the crests of the arms of the tidal spiral shock eclipse the inner areas of a disk. The heating of gas at the shock fronts should increase the vertical thickness of the regions of the disk that are seen at these orbital phases, providing an explanation for the observed brightness dips.

The transport of angular momentum plays an important role in the formation of the accretion disks of cataclysmic variables. Two mechanisms have been suggested: the first acts in the case of a standard Shakura–Sunyaev accretion disk [20] and involves turbulent viscosity, while the second considers the direct dissipation of energy by tidal spiral waves in the disk [21], which form a global structure. Such a global spiral structure was first discovered via Doppler mapping of IP Peg in outburst [22].

Tidal spiral structure can also be detected via photometric observations of the orbital light curves, if some parts of the spirals protrude over the surface of the disk. Such spiral protrusions were found by Hachisu *et al.* [23] in the accretion disk of V1494 Aql (Nova Aql 1999-2) in their studies of the photometric variability of this system. Hachisu *et al.* [23] modeled the orbital light curve using a model in which the binary system consists of a main-sequence star that completely fills its Roche lobe and a white dwarf surrounded by a circular accretion disk. The system was assumed to have a circular orbit, and mutual heating of the components was taken into account. All the components—the white dwarf, red star, and accretion disk—emit as blackbodies at a local temperature that depends on the position of the emitting area. To reproduce the behavior of the brightness of the system out of eclipse, it was assumed that a two-armed spiral shock was located at the surface of the disk. The accretion disk is an axisymmetric structure with radius $R_{\rm disk} = \alpha R_{\rm L1}$, where $R_{\rm L1}$ is the effective radius of the white dwarf's critical lobe and the height of the emitting surface over the equatorial plane $h = \beta R_{\rm disk}(\rho/R_{\rm disk})^2$, where ρ is the distance to the center of mass of the white dwarf in the equatorial plane. To model the two-armed spiral structure in a selected area of an unperturbed accretion disk, its height h was multiplied by a factor

$$z_1 = \max\left(1, \frac{\xi_1}{\sqrt{(\rho/R_{\rm disk} - \exp(-\eta(\phi-\delta)))^2 + \varepsilon^2}}\right), \quad (1)$$

$$z_2 = \max\left(1, \frac{\xi_2}{\sqrt{(\rho/R_{\rm disk} - \exp(-\eta(\phi - \delta - \pi)))^2 + \varepsilon^2}}\right),$$

$$Z = z_{\rm height} = \max(z_1, z_2).$$

At the disk edge, $0.9 < \rho/R_{\rm disk} < 1.0$, this coefficient in the region of spiral arms varies as

$$Z = z_{\rm height} - 10(z_{\rm height} - 0.25) \times (\rho/R_{\rm disk} - 0.9). \quad (2)$$

The coefficients ξ_1, ξ_2 specify the amplitude of the spirals. When $\xi_1 = \xi_2$, the spirals are symmetric and have similar heights; η determines the inverse pitch angle of the logarithmic spiral; ϕ is the azimuth of the considered area at the disk surface in a left-handed coordinate system (counterclockwise motion) and δ is the position angle of the spirals, with both δ and ϕ measured from the line connecting the binary components; ε defines the width of the spirals and, together with ξ_1, ξ_2, represents the height of the spirals at the edge of the disk for $\rho = R_{\rm disk}$, $\phi = \delta$: $z_{\rm height} \sim \max(\xi_1, \xi_2)/\varepsilon$.

Unfortunately, Hachisu *et al.* [23] do not provide any additional details of their model.

The aim of the present paper is to modify the Hachisu–Kato model [23] for cataclysmic binaries that have an elliptical disk surrounding a white dwarf and a hot line in the stream-disk interaction region, in order to describe the tidal spiral waves in the outer parts of the accretion disk.

2. MODEL OF THE SYSTEM

Modeling of the "Unperturbed" Inner Surface of the Disk

Following the algorithm used in our previous work, we will consider a close binary containing a Roche lobe filling, evolved, late-type star and a white dwarf surrounded by an elliptical accretion disk whose surface will be superimposed by spiral structures. The dwarf is represented by a sphere with radius R_w, and the shape of the red dwarf's surface is described by a Roche potential with filling factor $\mu = 1.0$. More details on the calculation of the shape of the Roche lobe are given in [24].

Our code takes into account the presence of the hot line in the region where the disk interacts with the stream of matter flowing from the red dwarf through the inner Lagrangian point. A detailed description of the algorithm used to model the hot line is presented in [25].

Three-dimensional gas-dynamical simulations of the gas flows in close binaries [17] have shown that the contribution of the hot line to the total radiation flux of the system is negligible compared to the flux from the accretion disk in the active state of a cataclysmic binary. This conclusion was confirmed by fitting light curves of the cataclysmic binary IP Peg in outburst carried out for both hot-line and hot-spot models [26]. This analysis showed that the hot-line model is virtually identical to the hot-spot model when describing the active state of a high-luminosity accretion disk whose radius is close to the radius of the white dwarf's inner Roche lobe, since the hot line degenerates into a small bulge (prominence) at the lateral surface of the disk in this case, and the parameters of the hot-line emission region are close to those of the hot spot in the corresponding model.

However, spiral shocks at the surface of an accretion disk are not observed only in the active states of close binaries. The same gas-dynamical simulations by Kuznetsov *et al.* [17] show that the spiral shocks do not disappear in quiescence, when the hot-line model is more justified as a description of the emission of the disk and circumdisk structures. To make our algorithm for modeling spiral perturbations at the surface of an accretion disk valid for both outburst and quiescent states of a cataclysmic binary, we applied the more complex close-binary model with a hot line, although this increases the computation time necessary for the light-curve modeling.

We described the shape of the inner surface of an elliptical disk in the $\overline{OXYZ}$ coordinate frame, in which the $\overline{OX}$ and $\overline{OY}$ axes are located in the orbital plane of the system and the $\overline{OZ}$ axis is perpendicular to this plane (Fig. 1). The white dwarf is located at a focus of the ellipsoid; the origin of the $\overline{OXYZ}$ frame is at this focus, the $\overline{OX}$ axis makes an angle α_e with the line connecting the system components and is directed toward the periastron of the disk ellipse, and the $\overline{OY}$ axis is perpendicular to the $\overline{OX}$ axis and is directed in the clockwise direction from the latter. The algorithm for calculating the shape of an "unperturbed" elliptical disk and its emission is described in detail in [3].

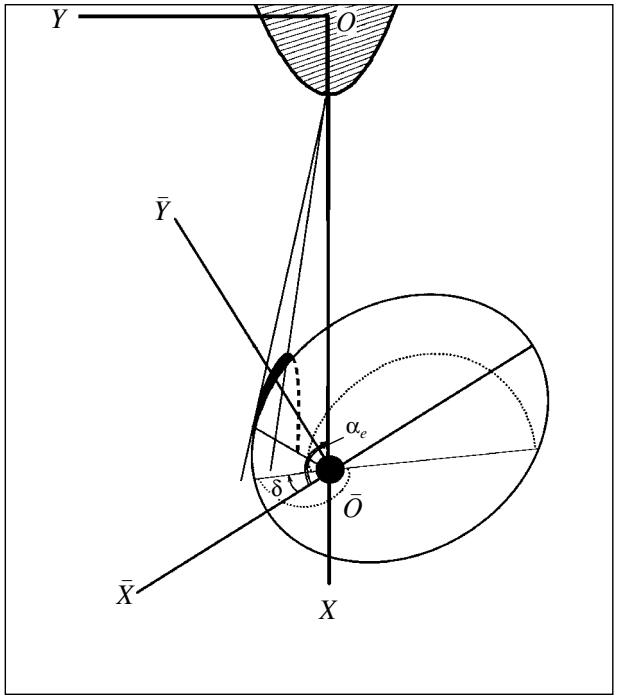

Fig. 1. Coordinate frames used to compute the spiral structure of the disk. The dashed lines show schematically the direction of the crests of the spiral arms. The angle between the spiral arms is 180°; α_e is the azimuth of the disk periastron. The position of the hot line, which is part of an ellipsoid with its center inside the disk and with one of its lateral surfaces coincident with the tangent extended to the disk from the inner Lagrangian point, is shown schematically. During outburst, nearly the entire hot-line ellipsoid is immersed in the disk. The small fraction of this ellipsoid extending above the disk surface (shaded) can be considered an analog of a hot spot.

The accretion disk around the white dwarf is defined as a complex figure corresponding to the intersection of an ellipsoid with semiaxes a, b, and c (the lateral surface of the disk) and a paraboloid (representing the inner surface of the disk). The paraboloid is defined by the parameter A_p, whose value depends on the rotation angle of the radius vector $\overline{\psi}$:

$$A_p(\overline{\psi}) = \frac{Ab^2}{a^2(1+e\cos\overline{\psi})} = \frac{A(1-e^2)}{(1+e\cos\overline{\psi})}. \quad (3)$$

Here, e is the eccentricity of the elliptical disk and A is a constant that is similar to the paraboloid constant introduced in [27] to describe the shape of the inner disk surface. In the case of a circular disk, for $e = 0$, $A_p = A$; in the usual case, $A_p = A(1-e)$ at periastron ($\overline{\psi} = 0.0$) and $A_p = A(1+e)$ at apoastron ($\overline{\psi} = \pi$). Here, $\overline{\psi}$ is the angle in the orbital plane between the positive $\overline{OX}$ axis and the projection of the radius vector $\overline{R}_i$ extending from the vertex of the parabola to a given point on the surface of the disk onto the orbital plane ρ_i (not the point $\overline{O}$, but a point located below the orbital plane a distance

$$z_0 = \frac{R_w^2}{A^2(1-e)^2}, \quad (4)$$

where R_w is the radius of the spherical white dwarf).

The shape of the unperturbed (i.e., not deformed by spiral-like structures) inner paraboloidal surface of the disk is defined by the equation

$$\rho^2 = \overline{x}^2 + \overline{y}^2 = A_p^2(\overline{\psi})\left(\overline{z} + \frac{R_w^2}{A^2(1-e)^2}\right). \quad (5)$$

Consequently, the equation of the paraboloidal surface is [3]

$$\Omega = (\overline{x}^2 + \overline{y}^2)\frac{(1+e\cos\overline{\psi}(\overline{x},\overline{y}))^2}{A^2(1-e^2)^2} - \overline{z} - \frac{R_w^2}{A^2(1-e)^2}, \quad (6)$$

and the components of the vector normal to this surface are

$$\overline{n}_x = -\frac{\Omega'_x}{|\Omega'|}, \quad \overline{n}_y = -\frac{\Omega'_y}{|\Omega'|}, \quad \overline{n}_z = -\frac{\Omega'_z}{|\Omega'|}, \quad (7)$$

where

$$\Omega'_x = \frac{2\rho(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}(\cos\overline{\psi} + e), \quad (8)$$

$$\Omega'_y = \frac{2(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}\overline{y},$$

$$\Omega'_z = -1,$$

$$|\Omega'| = \sqrt{(\Omega'_x)^2 + (\Omega'_y)^2 + (\Omega'_z)^2}.$$

An area element of the paraboloidal disk is defined by the formula derived in [3]:

$$dS_{\text{par}} = \frac{2\rho(\rho^2 + A_p^4(\overline{\psi}))\sqrt{\rho^2(1+e^2+2e\cos\overline{\psi}) + G_0^{-2}}\,d\overline{\eta}_d d\overline{\psi}}{A_p^4(\overline{\psi})(1+e\cos\overline{\psi})}, \quad \text{where} \quad G_0 = \frac{2(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}. \quad (9)$$

Modeling of a Spiral Shock

We will use the expression presented in [23] to define the positions of the spirals at the inner surfaces of the disk (the "upper" if $\overline{z} \geq 0$ and the "lower" if $\overline{z} < 0$), where it was assumed when describing the shape of the spirals at the surface of a flat, circular disk that the z coordinate of a point on the surface outside the spirals is equal to the disk height h, $z = h$, while the z coordinates of points inside spirals increase in accordance with (1) and (2).

We modified this expression as follows in order to model a spiral (or, in general, two spirals) in an elliptical disk with the complex form described above. We assume that, when $\rho < 0.9\,R_d$, the z coordinate of a point at the inner surface of the disk inside a spiral arm varies as

$$Z = z_s = zh \tag{10}$$
$$= z\frac{\xi}{\sqrt{[\rho/R_d(\overline{\psi}) - \exp(\eta(\overline{\psi} - \delta - B))]^2 + \varepsilon^2}}.$$

This takes into account the fact that, in our model, the position angle $\overline{\psi}$ is measured clockwise from the disk periastron, while the parameter B is different for different spiral arms, as in [23]: for the arm that is formed close to the periastron of the elliptical disk, $B = 0$ if $\overline{\psi} < \delta$ and $B = 2\pi$ if $\overline{\psi} > \pi + \delta$, where δ is the angle between the disk periastron and the radius vector at which the amplitude of the spiral has a maximum (δ is measured from the $\overline{OX}$ axis in the direction of the $\overline{OY}$ axis). If two spiral arms are present in the disk, we shall assume that the second has its maximum height in the region that is located symmetrically with respect to the first arm relative to the major axis of the elliptical disk, i.e., $B = \pi$, while the second arm itself is confined to the position angles $\pi + \delta > \overline{\psi} > \delta$. This specifies counterclockwise rotation of the spirals and does not allow them to have the same angle $\overline{\psi}$. Finally, according to Kuznetsov *et al.* [17], the amplitude of spiral-like structures rapidly decays in the inner regions of the disk. The amplitude of the logarithmic spiral very rapidly decreases as it approaches the white dwarf. Note that the radius of an elliptical disk R_d varies with the position angle of a point at the disk surface $\overline{\psi}$ as

$$R_d(\overline{\psi}) = \frac{a(1-e^2)}{1 + e\cos\overline{\psi}}. \tag{11}$$

As in [23], we shall limit the height of the spiral at the edge of the disk. We will assume that, at the very edge of the disk, when $0.9 < \rho/R_d \leq 1.0$, the coefficient Z decreases as

$$Z = z_s - z_{\text{edge}} \tag{12}$$
$$= z_s - 10(z_s - z_{\text{crit}})(\rho/R_d - 0.9).$$

In this case, for $\rho/R_d = 1$, we obtain $Z = z_{\text{crit}}$, where z_{crit} is the height of the unperturbed disk above the orbital plane [3]; for $\rho/R_d = 0.9$, we obtain $Z = z_s$; i.e., the height of an area element at the spiral is equal to the z coordinate given by (10).

The components of the vector normal to the surface of the spiral, specified by the potential

$$\Omega = (\overline{x}^2 + \overline{y}^2)\frac{(1 + e\cos\overline{\psi}(\overline{x},\overline{y}))^2}{A^2(1-e^2)^2} - \frac{R_w^2}{A^2(1-e)^2} \tag{13}$$
$$- z\xi\left[\left(\frac{\rho}{R_d} - \exp(\eta(\overline{\psi} - \delta - B))\right)^2 + \varepsilon^2\right]^{-0.5},$$

are

$$\overline{n}_{x,\text{sp}} = -\frac{\Omega'_{x,\text{sp}}}{|\Omega'_{\text{sp}}|},$$
$$\overline{n}_{y,\text{sp}} = -\frac{\Omega'_{y,\text{sp}}}{|\Omega'_{\text{sp}}|},$$
$$\overline{n}_{z,\text{sp}} = -\frac{\Omega'_{z,\text{sp}}}{|\Omega'_{\text{sp}}|},$$

where

$$\Omega'_{x,\text{sp}} = \frac{2\rho(1 + e\cos\overline{\psi})}{A^2(1-e^2)^2}(\cos\overline{\psi} + e) \tag{14}$$

$$- \frac{z\xi\left(\frac{\rho}{R_d} - \exp(\eta(\overline{\psi} - \delta - B))\right)\left(\frac{e + \cos\overline{\psi}}{a(1-e^2)} + \frac{\eta}{\rho}\sin\overline{\psi}\,\exp(\eta(\overline{\psi} - \delta - B))\right)}{\left[\left(\frac{\rho}{R_d} - \exp(\eta(\overline{\psi} - \delta - B))\right)^2 + \varepsilon^2\right]^{3/2}},$$

$$\Omega'_{y,\mathrm{sp}} = \frac{2\rho(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}\sin\overline{\psi}$$
$$- \frac{z\xi\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)\left(\dfrac{\sin\overline{\psi}}{a(1-e^2)} - \dfrac{\eta}{\rho}\cos\overline{\psi}\exp(\eta(\overline{\psi}-\delta-B))\right)}{\left[\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)^2 + \varepsilon^2\right]^{3/2}},$$

$$\Omega'_{z,\mathrm{sp}} = \frac{-\xi}{\sqrt{\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)^2 + \varepsilon^2}},$$

$$|\Omega'_{\mathrm{sp}}| = \sqrt{(\Omega'_{x,\mathrm{sp}})^2 + (\Omega'_{y,\mathrm{sp}})^2 + (\Omega'_{z,\mathrm{sp}})^2}.$$

At the edge of the disk, when $0.9 < \rho/R_d \le 1.0$, the derivatives of the potential (13) are

$$\Omega'_{x,\mathrm{sp}} = \frac{2\rho(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}(\cos\overline{\psi}+e) - 10(z_s - z_{\mathrm{crit}})\frac{e+\cos\overline{\psi}}{a(1-e^2)} - 10\left(1-\frac{\rho}{R_d}\right) \qquad (15)$$
$$\times \frac{z\xi\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)\left(\dfrac{e+\cos\overline{\psi}}{a(1-e^2)} + \dfrac{\eta}{\rho}\sin\overline{\psi}\exp(\eta(\overline{\psi}-\delta-B))\right)}{\left[\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)^2 + \varepsilon^2\right]^{3/2}},$$

$$\Omega'_{y,\mathrm{sp}} = \frac{2\rho(1+e\cos\overline{\psi})}{A^2(1-e^2)^2}\sin\overline{\psi} - 10(z_s - z_{\mathrm{crit}})\frac{\sin\overline{\psi}}{a(1-e^2)} - 10\left(1-\frac{\rho}{R_d}\right)$$
$$\times \frac{z\xi\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)\left(\dfrac{\sin\overline{\psi}}{a(1-e^2)} - \dfrac{\eta}{\rho}\cos\overline{\psi}\exp(\eta(\overline{\psi}-\delta-B))\right)}{\left[\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)^2 + \varepsilon^2\right]^{3/2}},$$

$$\Omega'_{z,\mathrm{sp}} = \frac{-10\xi\left(1-\dfrac{\rho}{R_d}\right)}{\sqrt{\left(\dfrac{\rho}{R_d} - \exp(\eta(\overline{\psi}-\delta-B))\right)^2 + \varepsilon^2}}.$$

In the region of a spiral, the area of a surface-area element increases as

$$dS_{\mathrm{spir}} \cong dS_{\mathrm{par}}/|\cos\gamma_{\mathrm{spir}}|, \qquad (16)$$

where $\cos\gamma_{\mathrm{spir}} = \overline{n}_x\overline{n}_{x,\mathrm{sp}} + \overline{n}_y\overline{n}_{y,\mathrm{sp}} + \overline{n}_z\overline{n}_{z,\mathrm{sp}}$ is the cosine of the angle between the normals to the point $(\overline{n}_x, \overline{n}_y, \overline{n}_z)$ at the paraboloid and to the corresponding point $(\overline{n}_{x,\mathrm{sp}}, \overline{n}_{y,\mathrm{sp}}, \overline{n}_{z,\mathrm{sp}})$ in the spiral, and dS_{par} is a unit area at the unperturbed paraboloidal surface of the disk [see Eq. (9)]. The accuracy of this expression increases with the number of unit areas at the disk surface.

3. VERIFICATION OF CONDITIONS FOR THE ECLIPSE AND HEATING OF THE DISK BY EMISSION FROM THE SYSTEM COMPONENTS

Heating of the disk surface. The method used to calculate the temperature of an area element on the paraboloidal part of the disk surface is described in detail in our earlier papers [3, 25, 27]. The heating is due primarily to the release of the gravitational energy of the matter and its conversion into heat as the matter approaches the surface of the white dwarf. It is usually assumed that the variation of this "basic" temperature follows the dependence [20]

$$T_g = T_w \left(\frac{R_w}{\rho}\right)^{\alpha_g}, \qquad (17)$$

where ρ is the distance from the white dwarf's center to the center of the area element on the disk in the orbital plane and α_g is a parameter. In quiescence, α_g is close to 0.75; i.e., it is assumed that every point on the disk surface radiates as a blackbody. In outburst, when the emission flux increases by a factor of hundreds, the value of α_g can be reduced to $\sim$0.1, as is shown by numerous observations and their analysis (see, e.g., [28, 29]). The heating of the disk surface by the radiation of the white and red dwarfs is less important. The method used to calculate the contribution of all these factors to the temperature of an area element is described in detail in [3]. The basic temperatures of areas located in the regions of the spiral arms increase by $T_{\rm spir}$, i.e., $T(i) = T_d(i) + T_{\rm spir}$.

Below, we consider a method for testing whether an area on the surface of a given component is shielded from irradiation (or eclipsed) by spiral arms that protrude above the surface of the unperturbed disk. This procedure will be applied again (for different radius vectors) to estimate the heating of areas on the disk by the radiation of the white and red dwarfs and to test whether the disk and white and red dwarfs are eclipsed by the spiral arms. Let us describe this procedure in its general form.

(i) Let us draw a ray $\overline{r}$ from the point $(\overline{x}_1, \overline{y}_1, \overline{z}_1)$ to the point $(\overline{x}_2, \overline{y}_2, \overline{z}_2)$. The first point may represent the coordinates of the observer or the center of an area (at the white dwarf or the pole of the red dwarf) that undergoes either irradiation or eclipse. The coordinates $(\overline{x}_2, \overline{y}_2, \overline{z}_2)$ correspond to the area under consideration, i.e., the area whose irradiation or eclipse we are testing. The length of this ray is then

$$|\overline{\mathbf{r}}_1| = \sqrt{(\overline{x}_1 - \overline{x}_2)^2 + (\overline{y}_1 - \overline{y}_2)^2 + (\overline{z}_1 - \overline{z}_2)^2}, \qquad (18)$$

and the coordinates of a unit vector along the ray are

$$\overline{\mathbf{n}}_r = \frac{\overline{r}_1}{|\overline{\mathbf{r}}_1|} = \left\{\frac{\overline{x}_1 - \overline{x}_2}{|\overline{\mathbf{r}_1}|}, \frac{\overline{y}_1 - \overline{y}_2}{|\overline{\mathbf{r}_1}|}, \frac{\overline{z}_1 - \overline{z}_2}{|\overline{\mathbf{r}_1}|}\right\}. \qquad (19)$$

(ii) Let us write the equation for the straight line passing along the vector through the center of the given area in the parametric form in the parameter τ

$$\overline{x} = \overline{x}_2 + \tau\overline{n}_{rx} = \overline{x}_2 + \tau\frac{\overline{x}_1 - \overline{x}_2}{|\overline{\mathbf{r}}_1|}, \qquad (20)$$

$$\overline{y} = \overline{y}_2 + \tau\overline{n}_{ry} = \overline{y}_2 + \tau\frac{\overline{y}_1 - \overline{y}_2}{|\overline{\mathbf{r}_1}|},$$

$$\overline{z} = \overline{z}_2 + \tau\overline{n}_{rz} = \overline{z}_2 + \tau\frac{\overline{z}_1 - \overline{z}_2}{|\overline{\mathbf{r}_1}|}$$

and test whether this line crosses the surface of the spiral arm between the point $(\overline{x}_1, \overline{y}_1, \overline{z}_1)$ and the trial area.

(iii) For this purpose, we must solve the system of equations for the straight line (20) and the equation specifying the shape of the disk with spiral arms simultaneously for τ, taking into account any supplementary conditions. The absence of solutions would indicate that the line of sight does not cross the body of the spiral arm; if the line of sight intersects the spiral arm, we assume that the area is shielded from irradiation. However, the equation for τ is rather complex. Therefore, we propose the following method. We specify for a given ray a step in τ, namely, $\tau = j\Delta\tau$. The size of the step and the number of steps can be different. Moreover, additional conditions can be used to terminate the testing.

(iv) Applying (20), we calculate the moving coordinates along the ray $(\overline{x}_j, \overline{y}_j, \overline{z}_j)$ and test whether a given point of the ray lies inside the region of the spiral arm. To do this, we insert the coordinates $(\overline{x}_j, \overline{y}_j)$ into the expression for the scaling coefficient h in the equation for paraboloidal disk with spiral arms:

$$h = \frac{\xi}{\sqrt{\left(\dfrac{\rho_j}{R_d(\overline{\psi})} - \exp(\eta(\overline{\psi} - \delta - B))\right)^2 + \varepsilon^2}}, \qquad (21)$$

where $\rho_j = \sqrt{\overline{x}_j^2 + \overline{y}_j^2}$. If $h \le 1$, the point under consideration is located outside the spiral-arm region.

(v) If $h > 0$, we calculate the height of the disk $\overline{z}_D$ for the given ray coordinates $(\overline{x}_j, \overline{y}_j)$ using the formula $z_D = \rho_j^2/A_p^2(j) - z_0$. Since A_p varies with the rotation angle of the radius vector $\overline{\psi}$ [see (3)], we must first find $\overline{\psi}$. We calculate $\overline{\psi_0} = \arcsin(\overline{y}_j/\rho_j)$, $\overline{\psi} = \overline{\psi}_0$ if $\overline{x}_j \ge 0$, $\overline{y}_j \ge 0$, $\overline{\psi} = \overline{\psi}_0 + 2\pi$ if $\overline{x}_j \ge 0$, $\overline{y}_j < 0$, and $\overline{\psi} = \pi - \overline{\psi}_0$ if $\overline{x}_j < 0$. The value of $\overline{z}_0$ is given by (4). If the inequality

$$\overline{z}_D h \le \overline{z}_2 + \tau_j\frac{\overline{z}_1 - \overline{z}_2}{|\overline{\mathbf{r}}_1|} \qquad (22)$$

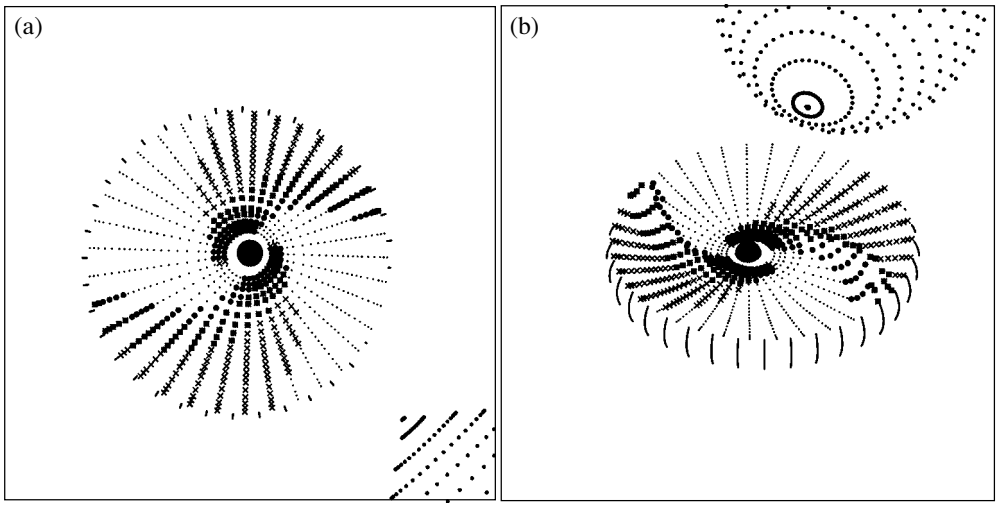

Fig. 2. Sketch of the areas at the inner part of the disk located behind the crests of the spiral arms and shielded from heating by the white dwarf. The sketch corresponds to the parameters $q = 1.7$, $i = 0°$ (a) and $50°$ (b), disk eccentricity $e = 0.10$, semimajor axis of the disk $a = 0.378a_0$, semithickness of the outer edge of the disk $z_{\text{crit}} = 0.0452a_0$, $R_w = 0.0277a_0$, and $\alpha_e = 85.0°$. The parameters of the spiral arms are $\xi = 0.2$, $\eta = 1.0$, $\varepsilon = 0.05$, and $\delta = 0°$. The crosses mark the areas shielded from heating by the crests of the spiral waves, the solid circles mark areas located in the spiral arms, and the crossed circles mark areas in the spirals that are shielded from heating by the crest of a spiral (the areas in front of the crest are susceptible to heating, while those behind the crest are shielded).

holds, the region of the disk with a spiral arm corresponding to the step j is located beneath the ray, and the ray does not intersect the disk surface. If this condition is not satisfied, we assume that the ray does intersect the disk, so that the area is either unseen (if the conditions for an eclipse are satisfied) or is shielded from irradiation (if the conditions for heating are satisfied).

(vi) If the moving point on the ray $\overline{\mathbf{r}}_1$ coincides with the edge of the disk, i.e., the relative value of the radius vector $\rho_j/R_d(\overline{\psi})$ is in the range 0.9–1.0, inequality (22) is replaced by the inequality

$$\overline{z}_D h - 10(\overline{z}_D h - \overline{z}_{\text{crit}})(\rho_j/R_d(\overline{\psi}) - 0.9) \leq \overline{z}_2 + \tau_j \frac{\overline{z}_1 - \overline{z}_2}{|\overline{\mathbf{r}}_1|}. \quad (23)$$

The test is continued until the ray touches the point $(\overline{x}_1, \overline{y}_1, \overline{z}_1)$ or until the supplementary conditions noted in item (iii) are satisfied.

Thus, to test the conditions for heating of an area on the disk surface, we replace the coordinates $(\overline{x}_2, \overline{y}_2, \overline{z}_2)$ with $(\overline{x}_d, \overline{y}_d, \overline{z}_d)$ and the coordinates $(\overline{x}_1, \overline{y}_1, \overline{z}_1)$ with the coordinates of the center of the area under study on the white dwarf $(\overline{x}_w, \overline{y}_w, \overline{z}_w)$. For the step in the parameter τ, we adopt $\Delta\tau = R_d(\overline{\psi})/100$ $(j = 1, \ldots, 100)$; the supplementary condition is the following: if inequalities (22) and (23) are satisfied, the intersection of the ray with the area is tested as long as expression $\overline{x}^2(j) + \overline{y}^2(j) > R_w^2$ is satisfied.

Figure 2a shows a sketch in the plane of the sky of the areas in the inner region of the disk that are shielded from heating by the white dwarf's radiation because they are located behind the crests of the spiral arms. The image is constructed for the system parameters $q = M_w/M_2 = 1.7$, $i = 0°$, disk eccentricity $e = 0.1$, semimajor axis of the disk $a = 0.378a_0$ (where a_0 is the distance between the centers of mass of components), $A = 2.8$ (the corresponding semithickness of the outer edge of the disk is $z_{\text{crit}} = 0.0452a_0$), $R_w = 0.0277a_0$, and $\alpha_e = 85.0°$; the parameters of the spiral arms are $\xi_1 = \xi_2 = 0.2$, $\eta = 1.0$, $\varepsilon = 0.05$, and $\delta = 0$. The crosses show areas shielded from irradiation by the crests of the spiral arms, solid circles show areas located in the region of spiral arms, and crossed circles show areas in the region of the spiral arms that are shielded from irradiation by the crests of the spiral (areas in front of the crests are susceptible to heating, while those behind the crests are shielded). Figure 2b shows the same system, but for the orbital inclination $i = 50°$. The crests of the spiral arms and shielding of the areas behind the crests from the white-dwarf radiation are clearly visible.

Heating of the paraboloidal part of the disk by the radiation of the polar regions of the optical star is insignificant compared to the two previous heating factors. Since the radiation of the red dwarf is weak compared to that of its companion and the radiation of regions of the red dwarf adjacent to the orbital plane is shielded by the edge of the disk, we will assume that an area on the disk can be heated by

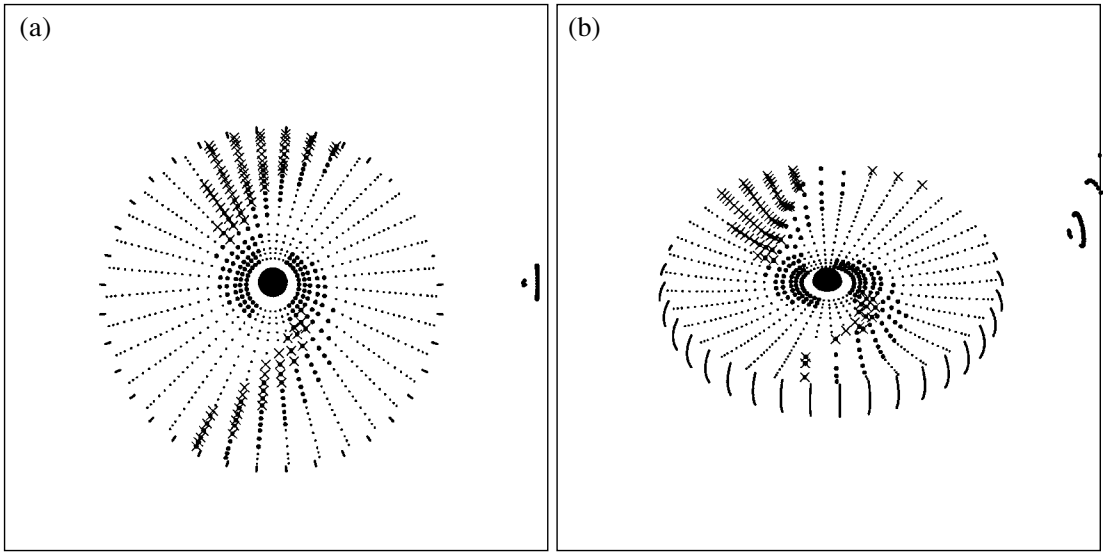

Fig. 3. The surface of the disk heated by the radiation of the red dwarf. The system parameters are the same as those used for Fig. 2. The orbital inclinations are $i = 0°$ (a) and 50° (b). The solid circles show areas located in the regions of spiral arms, and crosses show areas shielded from heating by the crests of the wave. We can clearly see that areas in front of the spiral arms crests are susceptible to heating, while those behind the crests are shielded.

a point source with coordinates $(\cos\alpha_e, \sin\alpha_e, \overline{z}_{\rm pol})$ and luminosity $L_{\rm pol}$ if (a) it is located on the opposite side of the paraboloidal disk from the red star, (b) it is not shielded from irradiation by the edge of the disk, (c) it is not shielded from irradiation by the white dwarf, and (d) it is not shielded from irradiation by the crests of each of the spiral arms. The test for the first three conditions is described in detail in [3]. Testing of the last condition following the procedure (i)–(vi) described above is carried out for the coordinates of the first and second points, $(\cos\alpha_e, \sin\alpha_e, \overline{z}_{\rm pol})$ and $(\overline{x}_d, \overline{y}_d, \overline{z}_d)$, respectively.

Figure 3 shows the surface of the disk heated by the red dwarf; the parameters of the system are as in Fig. 2, and the orbital inclinations are $i = 0°$ (Fig. 3a) and 50° (Fig. 3b). Also shown by the solid circles are areas that are located in the region of the spiral arms, and by crosses areas shielded from heating by the crests of the arms. We can clearly see that areas in front of the crests are susceptible to heating, while those behind the crests are shielded.

Figure 4 shows the dependence of the heights of areas on the disk on their position angle, i.e., the z coordinates of the areas (upper panel) and their local temperatures, obtained by summing all the heating factors (lower panel); the parameters of the system are the same as those used for the sketch in Fig. 2. The temperature of the white dwarf was assumed to be the standard value, $T_w = 20\,000$ K, and the additional temperature due to the heating of matter by the spiral shock was taken to be $T_{\rm spir} = 500$ K. The figure clearly shows that the largest temperature variations are displayed by area elements in the inner part of the disk, close to the surface, and that they are due mainly to variations of the distance between the element and white dwarf due to the ellipticity of the disk. The effects of shielding of area elements from heating by the spiral-arm crests are more appreciable in the outer regions of the disk (lower panel of Fig. 4); Fig. 4 clearly shows the drop in the temperatures of area elements located behind the crest of the corresponding spiral arm.

Testing conditions for eclipse. The general principle underlying the testing of the visibilities of area elements on the surfaces of all the components is described in detail in our earlier papers. Briefly, it is as follows. Knowing the coordinates (x_i, y_i, z_i) of a given area on a selected component (which may be the red dwarf, white dwarf, hot-line ellipsoid, the paraboloidal or elliptical region of the disk, or the region of the spiral arms on the disk) in the coordinate frame $OXYZ$ fixed to the red dwarf (Fig. 1) and the components of the unit vector parallel to the line of sight,

$$\begin{aligned} o_x &= \cos\varphi_0 \sin i, \\ o_y &= \sin\varphi_0 \sin i, \\ o_z &= \cos i, \end{aligned} \tag{24}$$

where i is the orbital inclination of the system and φ_0 the orbital phase under consideration, we determine the angle γ between the normal vector to the

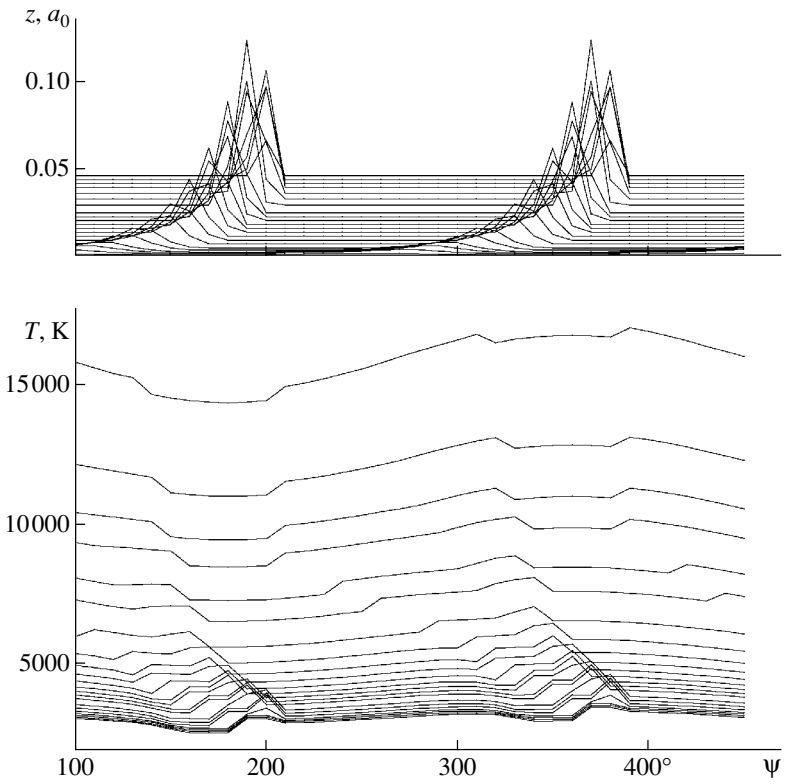

Fig. 4. Dependence of the z coordinates of the disk areas (upper panel) and their local temperatures (lower panel) on the position angle. The temperatures are obtained by summing all the heating factors. The parameters of the system are as in Fig. 2; $T_w = 20\,000$ K, $T_{\rm spir} = 500$ K.

area element $\mathbf{n}$ in the $OXYZ$ frame and a segment that is directed opposite to the line of sight:

$$\cos\gamma = \frac{\mathbf{n}\cdot\mathbf{o}}{|\mathbf{n}||\mathbf{o}|}. \tag{25}$$

The visibility condition for the area, i.e., the condition for the absence of self-eclipse, is $\cos\gamma > 0$. If this condition is satisfied, we check for the eclipse of the area element by other components of the system. For this, we construct a straight line in parametric form through the center of the area element parallel to the line of sight,

$$\begin{aligned} X &= x_i + to_x, \\ Y &= y_i + to_y, \\ Z &= z_i + to_z, \end{aligned} \tag{26}$$

and determine whether this line crosses the surfaces of other components located between the observer and the area element. As when testing for heating of area elements by radiation from the system components, in this case, we must simultaneously solve the system of equations (26) with the equation describing the shape of the component under consideration in the $OXYZ$ frame for the parameter t, taking into account any supplementary conditions. The absence of a solution indicates that the line of sight does not cross the body of a particular component. If the intersection occurs outside the region where the supplementary conditions are satisfied or beyond the region between the observer and the area element, we also assume that the area is not shielded and that we must include its contribution to the total flux, in accordance with its calculated temperature, size, and viewing angle γ. However, due to the complexity of the resulting expression for t, we will apply the shielding verification algorithm (i)–(vi) described above to test for shielding by the crests of spiral arms.

Testing for eclipse of the red dwarf by the spiral-arm crests. To determine the step in the parameter t, let us first find the coordinates of the point where the ray defined by t intersects the boundaries of the paraboloidal region of the disk. This can easily be accomplished during the testing for shielding of an area element on the star by the paraboloidal disk (for more details, see [3]). An area element on the red-dwarf surface has the coordinates (x, y, z) in the $OXYZ$ frame, while the coordinates of the boundary points of the paraboloidal disk are (x_{p1}, y_{p1}, z_{p1}) and (x_{p2}, y_{p2}, z_{p2}), respectively. Let us determine the length of the radius vectors drawn from the center of

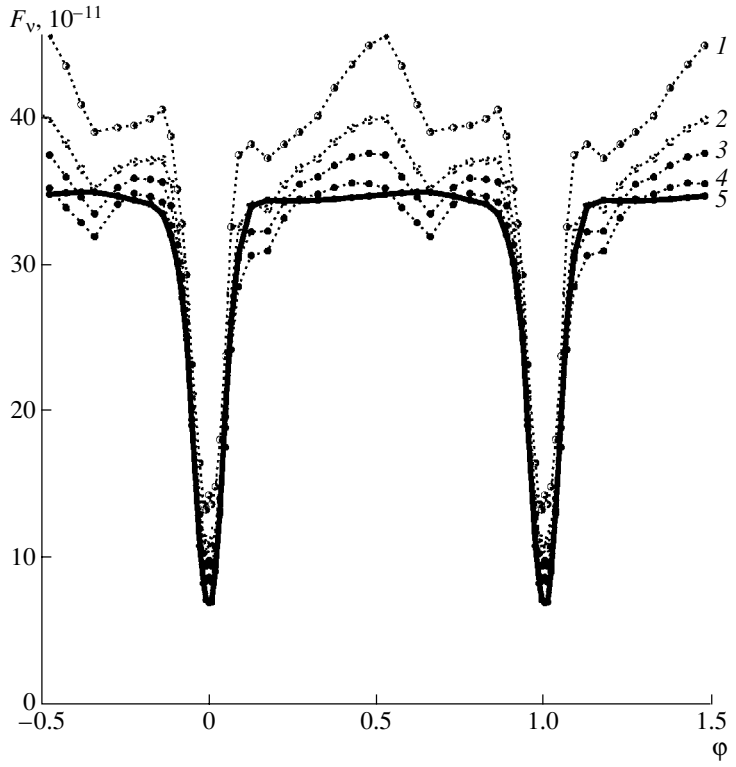

Fig. 5. Model curves representing the contribution of the accretion disk to the total emission from the system in the V filter as a function of the increase in the temperature in the spiral arm $T_{\rm spir}$ due to heating by the spiral shock (for $e = 0.1$, $\alpha_e = 120°$). Curves *1*, *2*, *3*, and *4* correspond to $T_{\rm spir} = 2000, 1000, 500$, and 0 K, respectively. The parameters of the spiral arms are $\xi_1 = \xi_2 = 0.15$, $\eta = 1.0$, $\varepsilon = 0.05$, and $\delta = 10°$; it is assumed that the angle between the two spiral arms is 180°. Curve *5* (bold) shows the input of the disk without the spiral arms).

the area element under consideration on the stellar surface to these points, applying formulas similar to (18), and find their difference Δr_{12}. In this way, we shall define the step in t to be $\Delta t = \Delta r_{12}/N$, where N is the selected number of steps. For our grid $(R, \overline{\eta}_d, \overline{\psi})$ defining the paraboloidal disk, $N = 100$ is sufficient.

Next, first applying formulas (26), we calculate the moving coordinates of the point belonging to the ray in the $OXYZ$ frame and, then, using the formulas

$$\begin{aligned} \overline{x} &= (1-x)\cos\alpha_e + y\sin\alpha_e, \\ \overline{y} &= (1-x)\sin\alpha_e - y\cos\alpha_e, \end{aligned} \qquad (27)$$

determine their values in the $\overline{OXYZ}$ frame; further, we apply steps (iv)–(vi) from the algorithm described above.

Testing for shielding of the white dwarf, hot-line ellipsoid, and paraboloidal disk by the crests of the spiral arms. In this test, we use the same step in t as in the algorithm used to test for heating of an area element on the paraboloidal disk by radiation from the white dwarf. As in the previous case, we first calculate the coordinates of a point belonging to the parametric curve in the $OXYZ$ frame and then use (27) to determine their values in the $\overline{OXYZ}$ frame and apply steps (iv)–(vi) of the algorithm above.

4. MODEL COMPUTATIONS

We tested the sensitivity of the code to various parameters, such as the temperature of the spiral arms $T_{\rm spir}$ and longitude of the periastron of the elliptical disk. Since spiral waves have thus far been observed in the disks only of systems with component-mass ratios $q \sim 1$–3, we used a system with parameters close to those of the dwarf nova IP Peg derived by us earlier [9] for the test runs: $q = M_w/M_2 = 1.7$, $i = 80°$, $T_{\rm eff} = 3400$ K, $R_w = 0.0061a_0$, and $T_w = 16\,000$ K. For the parameters of the disk, we used values close to those obtained by us in [26] during our study of a series of light curves of IP Peg in outburst. Namely, the radius of the disk at apoastron was taken to be $R_{\max} = 0.9\xi_w$, where ξ_w is the distance between the center of the white dwarf and the inner Lagrangian point L1; the parameter specifying the radial temperature distribution of the disk [Eq. (17)] was decreased to $\alpha_g = 0.575$ (the radial temperature distribution became flatter); the temperature of the

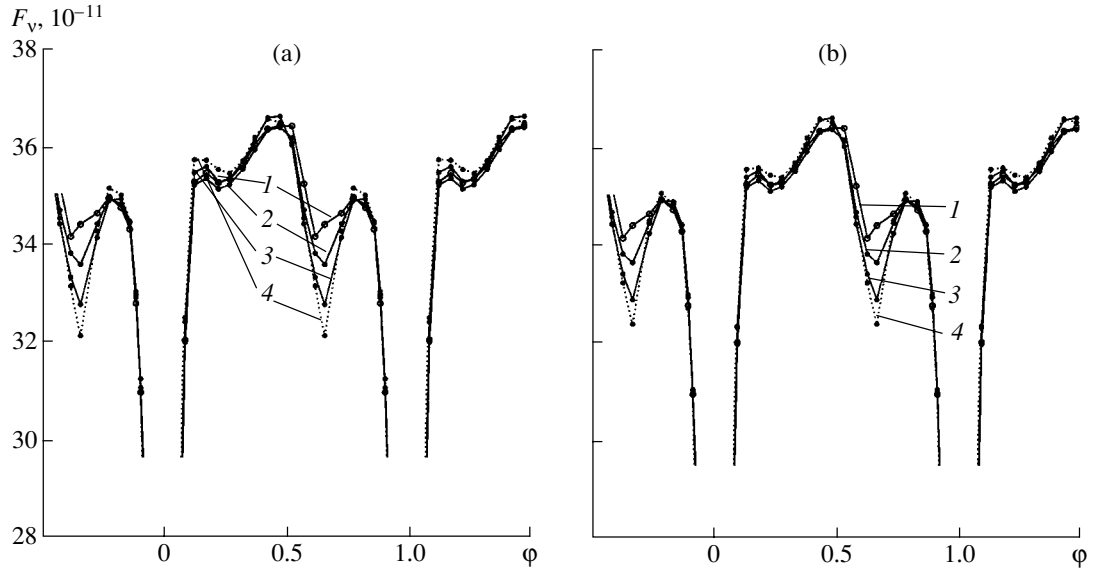

Fig. 6. Dependence of the out-of-eclipse contribution of the accretion disk to the total flux for $\xi = \xi_1 = \xi_2$ when two (a) and one (b) spiral arms are present at the surface of the disk. The curves *1*, *2*, *3*, and *4* are calculated for the same parameters of the system components as in Fig. 5 and for $\xi = 0.10$, 0.12, 0.14, and 0.18, respectively.

matter in the surface layers of the disk reached $T_{\rm in} = 72\,000$ K; and the height of the outer edge of the disk above the orbital plane was $z_{\rm crit} = 0.048a_0$—i.e., the disk was fairly thick (the corresponding value of the paraboloid constant in Eq. (3) is $A = 3$).

Figure 5 shows model curves representing the contribution of the accretion disk to the total emission from the system in the V filter as a function of $T_{\rm spir}$, the additional temperature increase in the spiral arm due to heating by the spiral shock. The disk ellipticity is $e = 0.1$, and the azimuth of the disk periastron is $\alpha_e = 120°$. Curves *1*, *2*, *3*, and *4* correspond to $T_{\rm spir} = 2000, 1000, 500$, and 0 K, respectively. Curve *5* (bold) shows the contribution of the disk without the spiral arms. Curves *1*–*4* were calculated using the parameters of the spiral arms $\xi_1 = \xi_2 = 0.15$, $\eta = 1.0$, $\varepsilon = 0.05$, and $\delta = 10°$, assuming that the angle between the spiral arms is 180°. The figure shows clearly that, even in the absence of additional heating of the matter in the spiral arms, curve *4* differs strongly from the almost symmetric curve representing the contribution of a disk unperturbed by spiral shocks (curve *5*); the reason for this is shielding of hot area elements in the inner regions of the disk by the crests of the spiral arms. The spiral arms contribute to the total emission from the disk both during the eclipse by the red dwarf and in out-of-eclipse orbital phases. Considerable distortions of the flux are seen out of eclipse, which become more prominent with increasing $T_{\rm spir}$ and are manifested as dips at orbital phases $\varphi \sim 0.2$ and $\sim$0.7. The phase of the dip minima does not depend on $T_{\rm spir}$, due to the similar azimuths of the spiral arms ($\delta = 10°$ and $\delta + B = 190°$) assumed in the model.

Another parameter of the spiral shock that strongly influences the shape of the light curve is ξ. This parameter specifies the height of the spiral arm above the surface of the paraboloidal disk in its outer regions [see comments about Eq. (2)]. The dependence of the out-of-eclipse contribution of the disk to the total flux when $\xi = \xi_1 = \xi_2$ is shown in Figs. 6a and 6b for cases when either two or one spiral shock is present on the disk surface. Curves *1*, *2*, *3*, and *4* were calculated for the same system parameters as were used for Fig. 5 and for $\xi = 0.10$, 0.12, 0.14, and 0.18, respectively. When ξ increases from $\sim$0.1 to $\sim$0.2, the amplitude of the dip in the light curve of the disk contribution at orbital phases $\varphi \sim 0.7$–0.8 also grows, in both the two-arm and one-arm cases. The influence of ξ on the shape of the light curve at phases $\varphi \sim 0.2$–0.25 is considerably weaker: the difference remains appreciable for the two-armed spiral, but the flux from the disk at this phase for the one-armed spiral is virtually independent of ξ.

Figure 7a shows the dependence of the contribution of the disk emission to the total flux on the azimuth of the disk periastron α_e for a two-armed spiral. The curves *1*, *2*, *3*, *4*, and *5* were calculated for the same parameters as for the curves discussed above and for $\alpha_e = 60°$, 80°, 100°, 120°, and 140°, respectively. The apoastron radius of the disk is $R_{\rm max} = 0.9\xi_w = 0.4989a_0$, the periastron radius is $R_{\rm min} = 0.4082a_0$, and the disk eccentricity is $e = 0.1$; it was assumed that $T_{\rm spir} = 2000$ K. For small values of the periastron azimuth (curves *1* and *2*), the second dip at

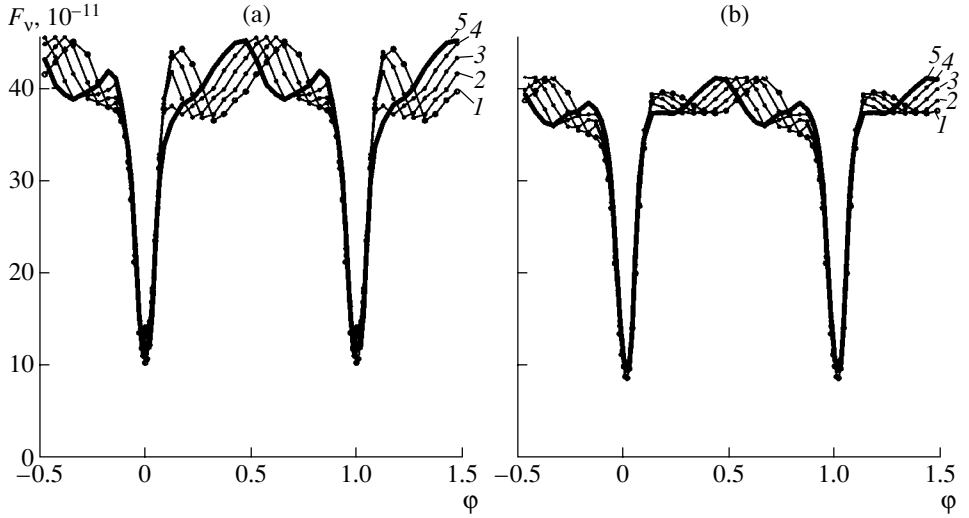

Fig. 7. Dependence of the contribution of the disk emission to the total flux from the system on the azimuth of the periastron for a two-armed spiral shock (a) and a one-armed shock located near the disk periastron (b). Curves *1*, *2*, *3*, *4*, and *5* correspond to $\alpha_e = 60°$, $80°$, $100°$, $120°$, and $140°$, respectively. The apoastron radius of the disk is $R_{\max} = 0.9\xi_w = 0.4989a_0$, the periastron radius is $R_{\min} = 0.4082a_0$, $e = 0.1$, and $T_{\rm spir} = 2000$ K. See the text for the remaining parameters.

phases $\varphi > 0.7$ nearly disappears, as it merges with the decline of the disk flux in the eclipse. The dip at orbital phases $\varphi \sim 0.2$–0.25 is the only significant feature of the disk emission curve. On the contrary, for large values of α_e (curves *4* and *5*), the dip at phases $\varphi \sim 0.7$ is clearly visible, while the dip at phases $\varphi \sim 0.2$ is smeared and resembles a variation of the slope of the emission during the egress from the eclipse. The depth and shape of the disk eclipse also vary with the periastron azimuth. Note that the amplitude of the magnitude drop at phases $\varphi \sim 0.2$ and $\varphi \sim 0.7$ is determined, not by the influence of a particular arm, but by the common influence of both spiral arms at the disk surface. This follows from Fig. 7b, which shows the disk contribution to the total flux assuming the same parameters as for Fig. 7a, but with a one-armed spiral located near the disk periastron. We can clearly see that the curve for this contribution changed only weakly, and the amplitude of the out-of-eclipse variations diminished.

5. APPLICATION OF THE SPIRAL-WAVE MODEL TO DETERMINE THE PARAMETERS OF IP Peg IN OUTBURST

We used our observations of IP Peg obtained during a "normal" outburst of this system [26] to test the suitability of our code for analyses of real cataclysmic variables. IP Peg is a well-studied cataclysmic variable with a so-called "double eclipse" in the light curve in its quiescent state. Such an eclipse is observed for a very small number of cataclysmic variables with large orbital inclinations, and it can be used to follow in detail the ingress and egress of various components of the system—the white dwarf, disk, and region of energy release by the edge of the disk. The parameters of IP Peg and a history of studies of this system are described in detail in [9], which is concerned with determining the parameters of this system in the quiescent state in the hot-line model. The orbital period of IP Peg is $P_{\rm orb} = 3.8$ h, the outbursts occur at intervals of $\sim$100 days, and the brightness of the system increases by $\sim 2.5^m$ in outburst, with the out-of-eclipse V brightness increasing from $\sim 14.7^m$ to $\sim 12.3^m$. The radius of the disk experiences a significant growth in outburst [30]: it is $\sim 0.24a_0$ in the quiescent state and increases to $\sim(0.34$–$0.37)a_0$ in the active state, increasing further with the system brightness, while the radial temperature profile deviates significantly from the classical "3/4" dependence. High-accuracy spectral observations of IP Peg in outburst show the presence of a large-scale, bright azimuthal structure in the outer regions of the disk [22, 31–35]. These structures resemble spiral arms in Doppler tomograms, with one of the arms located in the disk region near the extension of the gas-stream axis and the second arm located symmetrically relative to the first and displaced in azimuth by $\sim 180°$. The contribution of the spiral arms to the total flux is 12–15% of the disk flux, while the intensity of the arm emission can differ by a factor of a few units [13].

The V light curve of IP Peg (Table 1) obtained using a CCD array at the Crimean station of the Sternberg Astronomical Institute during an outburst on October 30, 2000 (JD 2451848) is shown in Fig. 8

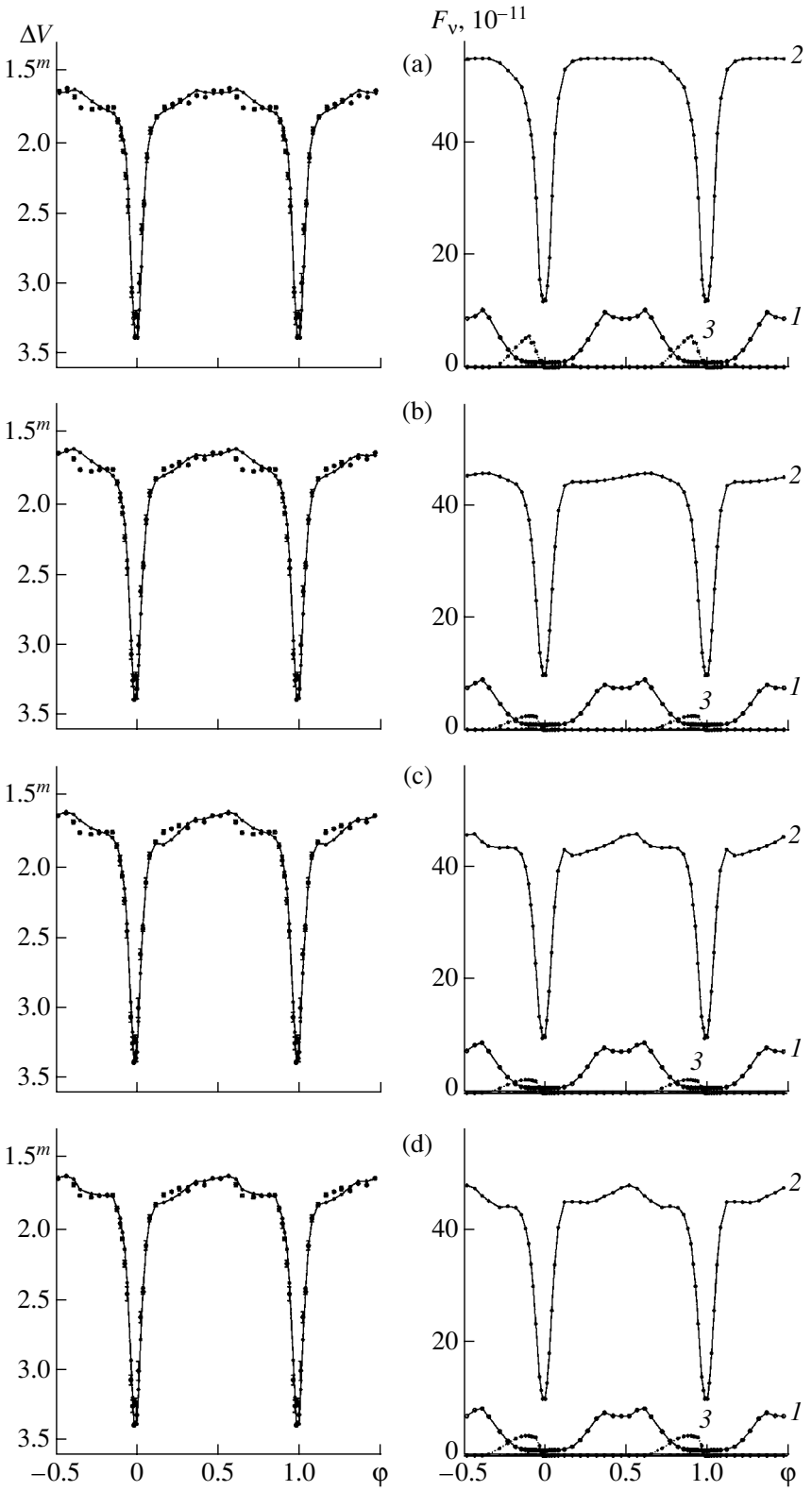

Fig. 8. Results of the fitting the light curve of IP Peg for the October 30, 2000, outburst using four models. The left column of plots shows the observed light curve (points with error bars) and synthesized light curves (solid curves) for the parameters listed in Table 2. The four plots correspond to the hot-spot (a) and hot-line (b) models, and models with two-armed (c) and one-armed (d) spiral shocks at the surface of the accretion disk. The plots to the right show the contribution to the total flux of the system by its components: the red dwarf *1*, accretion disk *2*, and region of energy release *3*. The contribution of the white dwarf is not shown since it is negligibly small due to the small size of the star and remains constant during the eclipse.

by points with their error bars. The presence of a "step" with a depth of up to $\sim 0.2^m$ at orbital phases 0.67–0.70, which subsequently increases by $\sim 0.05^m$ at phases 0.80–0.85, is characteristic of both this light curve and other orbital light curves obtained during this outburst. These outburst light curves do not have the orbital humps that are usual for the light curves in quiescence, and the flux at ingress does not differ much from the flux at egress.

We fit this light curve using hot-spot (the code described in [27]) and hot-line [25] models, as well as in models in which we applied the code described above to allow for the presence at the surface of the accretion disk of either a one-armed spiral shock located near the periastron of the elliptical disk or a two-armed spiral shock, with the second arm located near the disk apoastron. We applied the Nelder–Mead method [36] to search for the best-fit parameters (this algorithm is described in detail in our earlier papers; see, e.g., [29]). The results of the fits obtained for these four models are presented in Table 2; the light curves synthesized with the best-fit parameters are shown by the solid curves in Fig. 8. The mean error of the fitted parameters is 10–12%. Some of the parameters (q, i, $T_{\rm eff}$, R_w, η, and ε) were kept fixed, while others (T_w, ξ, $R_{\max}$) were varied within a narrow range of allowed values; this reduced the number of parameters to be searched for.

All models describe the shape and depth of the eclipse with sufficient accuracy. In all the models, the size of the disk is close to the size of the Roche lobe of the white dwarf, and the dependence of the temperature on the disk radius is fairly flat and corresponds to $\alpha_g \sim 0.544$–0.545 [see Eq. (17)]. The main differences between the models are manifested in the out-of-eclipse part of the light curve.

In the commonly accepted model with a hot spot located at the lateral surface of a circular disk (Fig. 8a″ left plot), the additional temperature increase in the hot-spot region is nearly equal to zero (as indicated by the fitting solution). In other words, there is no hot spot. This is expected, since the wings of the eclipse light curve are symmetric due to the absence of an orbital hump at phases $\varphi \sim 0.7$–0.8. The shape of the out-of-eclipse part of the synthesized light curve is essentially symmetric, and it is not able to describe the observed stepwise behavior of the brightness. Accordingly, the χ^2 residuals ($\chi^2 = 622$) are larger for this than for the other models. Below, we list the contribution of the system components to the total flux in arbitrary units. The maximum of the contribution (per unit wavelength) of the red dwarf is $10.3 \times 10^{-11} a_0^2$ erg s^{-1} cm^{-3}, of the disk is $55.3 \times 10^{-11} a_0^2$ erg s^{-1} cm^{-3}, and of the hot-spot region is $5.7 \times 10^{-11} a_0^2$ erg s^{-1} cm^{-3} (in fact, this is

Table 1. V Light curve of IP Peg during the outburst on JD 2451848

Orbital phase φ	ΔV	σ
0.0080	3.3043	0.0688
0.0186	2.9922	0.0676
0.0322	2.6116	0.0387
0.0481	2.4306	0.0195
0.0657	2.1038	0.0330
0.0878	1.9119	0.0241
0.1257	1.8148	0.0138
0.1748	1.7459	0.0142
0.2240	1.7230	0.0071
0.2696	1.7007	0.0150
0.3241	1.7165	0.0078
0.3729	1.6637	0.0072
0.4252	1.6757	0.0041
0.4762	1.6276	0.0027
0.5249	1.6297	0.0045
0.5741	1.6119	0.0054
0.6196	1.6719	0.0129
0.6593	1.7499	0.0100
0.7279	1.7606	0.0102
0.7770	1.7521	0.0047
0.8254	1.7445	0.0069
0.8608	1.7475	0.0109
0.8862	1.8418	0.0185
0.9063	1.9480	0.0352
0.9197	2.0580	0.0125
0.9337	2.2327	0.0238
0.9490	2.4477	0.0499
0.9724	3.0580	0.0359
0.9830	3.2430	0.0441
0.9913	3.3800	0.0046
0.9997	3.2150	0.0295

Table 2. Parameters of IP Peg in the V filter during the October 30, 2000, outburst in various models

Parameters*	Model			
	Hot spot	Hot line	Two-armed spiral	One-armed spiral
General parameters of the system				
$q = M_w/M_2$	1.71	1.71	1.71	1.71
i, deg	80.2	80.2	80.2	80.2
T_{eff}, K	3400	3400	3400	3380
$\langle R_2\rangle/a_0$	0.344	0.344	0.344	0.344
T_w, K	16 895	16 691	16 700	16 600
R_w/a_0	0.0061	0.0061	0.0061	0.0061
ξ_w/a_0	0.555	0.555	0.555	0.555
Disk parameters				
e	0.0	0.1251	0.1518	0.1524
$R_{\max}/\xi_w$	0.834	0.937	0.940	0.942
a/a_0	0.4629	0.4620	0.4528	0.4537
$R_{\min}/a_0$	0.4629	0.4041	0.3840	0.3846
$R_{\max}/a_0$	0.4629	0.5198	0.5215	0.5229
z_{crit}/a_0	0.0346	0.0411	0.0368	0.0428
α_e, deg	–	113.918	104.930	104.869
α_g	0.5490	0.5451	0.5436	0.5438
T_{in}, K	74 585	71 190	70 850	70 115
Parameters of hot line				
a_v/a_0	–	0.1293	0.1030	0.1036
b_v/a_0	–	0.156	0.152	0.150
c_v/a_0	–	0.0656	0.0565	0.0661
$T_{v(2)}$, K	–	1842	1659	1961
Parameters of hot spot				
r_{spot}/a_0	0.162	–	–	–
φ_{spot}	0.955	–	–	–
f_{spot}	0.00009	–	–	–
T_{spot}, K	6 953	–	–	–
Parameters of the spiral waves on the disk				
ξ	–	–	0.1318	0.1374
η	–	–	1.00	1.00
ε	–	–	0.05	0.05
δ, deg	–	–	9.36	9.27
T_{spir}, K	–	–	535.45	529.75
χ^2	622	529	425	330

* a_0 is the distance between the centers of mass of the white and red dwarfs; ξ_w is the distance from the center of mass of the white dwarf and the inner Lagrangian point L1; $T_{v(2)}$ is the temperature of the matter on the leeward side of the hot line (see the text); f_{spot} is a parameter used to estimate the temperature increase inside the hot spot relative to the temperature of the disk without the hot spot, $T_{\text{spot}} = T_d(1 + f_{\text{spot}})$.

the flux from the region at the lateral surface of the disk where the hot spot is formally located; because the emitting part of the disk is receding from this region, its contribution to the total flux at ingress is smoother that at egress). The contribution of the white-dwarf emission is constant out of eclipse and is only $0.15 \times 10^{-11} a_0^2$ erg s^{-1} cm^{-3} (this curve is not shown in the figure).

The out-of-eclipse part of the light curve is described somewhat better in the hot-line model (Fig. 8b). In the situation under consideration, the hot line is manifested as a small protuberance on the lateral side of the disk, which essentially does not differ from a hot spot. The region of the ellipsoid that is located on the side of the inflowing stream is completely immersed in the body of the disk and the pole of the ellipsoid is located near the lateral surface of the disk, so that only a slight increase of the emission flux from the leeward side of the hot line is observed (a sketch of this situation can be seen in Fig. 1, where the part of the hot line protruding over the lateral surface of the disk is shaded). Here, the region of energy release is quasi-ellipsoidal in shape, with the minor axis being comparable to the semithickness of the outer edge of the disk and the major axis being twice this size. In this model, the out-of-eclipse fraction of the synthesized light curve is already asymmetric, but this asymmetry is due primarily to the elliptical shape of the disk (Fig. 8b, right plot). The contribution of the hot-line emission to the total flux is negligible, as in the previous case. The decrease in the residuals is due primarily to the more accurate representation of the eclipse light curve.

Both the two-armed (Fig. 8c) and one-armed models (Fig. 8d) that take into account the presence of spiral shocks on the disk surface reproduce the variations of the out-of-eclipse brightness of IP Peg in outburst somewhat better than the previous models. In the case of a one-armed spiral (Fig. 8d), a step at phases $\varphi \sim 0.65$–0.85 clearly forms on the light curves synthesized for the best-fit parameters. The plot to the right shows that the formation of the step is determined purely by the emission of the disk with the one-armed spiral structure on its surface: for a given disk orientation, the crest of the spiral wave shields both the hotter parts of the disk that are close to the white dwarf and area elements that are heated by radiation from the white and red dwarfs. The step is less steep in the two-armed spiral model, due to the contribution of the supplementary emission of area elements located on the second arm and heated by the flux from the white dwarf at these phases (Fig. 2); moreover, deviations of the synthesized light curve from the observed light curve at phases $\varphi \sim 0.12$–0.25 are more substantial in this case and are due to the eclipse of the inner hot parts of the disk by the crest of the second spiral arm. This effect is absent in the one-armed model. The contribution of the spiral structure to the total flux is 9–18% for the one-armed spiral and up to 25–28% for the two-armed spiral.

In general, for system parameters that are close to the parameters of a real cataclysmic variable, namely, IP Peg, the influence of the spiral arm located near the disk apoastron on the light curve is not large and, due to the uncertainties in the photometric observations, it is difficult to distinguish between the presence of a one-armed or two-armed structure in the disk. During other outbursts of IP Peg, steps in the out-of-eclipse brightness were observed at both orbital phases $\varphi \sim 0.15$ and ~ 0.7 [32, 34, 37].

6. CONCLUSION

We suggested earlier a model that includes an extended region of energy release located along the gas stream flowing around an elliptical disk surrounding the white dwarf in a close binary system. This model enables us to describe the observed light curves of cataclysmic variables more fully than the usual hot-spot model and its modifications. The hot-line model is in good consistency with gas-dynamical models of the mass flows in semidetached close binary systems. These latter models should be used as the basic models for interpreting the light curves of cataclysmic binaries in quiescence and can provide a suitable explanation for the anomalous light curves of binary systems, the details of the "double eclipses" in systems with large orbital inclinations, and the light curves of U Gem-type systems, which have photometric minima related not to the eclipse of the white dwarf, but to the eclipse of the hot line in the vicinity of the white dwarf. For the outburst state of cataclysmic variables, gas-dynamical computations provide evidence for the appearance of an additional source of variability, related to the formation of a spiral structure in the outer regions of the accretion disk. We have presented here a light-curve synthesis algorithm for close binaries with an evolved late spectral type component and a white-dwarf component surrounded by an accretion disk possessing a spiral structure and with a region of energy release located outside the accretion disk. This makes it possible to reproduce both the light curve of the eclipse itself and the out-of-eclipse parts of the outburst light curves with sufficient accuracy. The set of algorithms we have devised enables us to represent more completely the variety of the light curves of cataclysmic variables, both in outburst and in quiescence, when an increase in the vertical thickness of the disk may occur in some close binaries, which can be considered a manifestation of a one-armed spiral structure in the disk.

ACKNOWLEDGMENTS

The author thanks A.M. Cherepashchuk for discussions and his continuing interest in the study and N.A. Katysheva and S.Yu. Shugarov for providing the observational data on IP Peg. This study was supported by the Russian Foundation for Basic Research (project nos. 02-02-17524 and 02-02-16462), the program "Leading Scientific Schools of Russia" (grant no. NSh-388.2003.2), the program "Russian Universities—Fundamental Studies" (project no. UR.02.03.012/1), and the Federal Science and Technology Program in Astronomy.

REFERENCES

1. V. G. Gorbatskii, Astrofiz. **3**, 245 (1967).
2. J. Smak, Acta Astron. **20**, 312 (1970).
3. T. S. Khruzina, Astron. Zh. **77**, 510 (2000) [Astron. Rep. **44**, 446 (2000)].
4. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **74**, 880 (1997) [Astron. Rep. **41**, 786 (1997)].
5. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **74**, 889 (1997) [Astron. Rep. **41**, 794 (1997)].
6. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).
7. M. Makita, K. Miyawaki, and T. Matsuda, Mon. Not. R. Astron. Soc. **316**, 906 (2000).
8. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **75**, 40 (1998) [Astron. Rep. **42**, 33 (1998)].
9. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **78**, 625 (2001) [Astron. Rep. **45**, 538 (2001)].
10. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **80**, 239 (2003) [Astron. Rep. **47**, 214 (2003)].
11. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **80**, 919 (2003) [Astron. Rep. **47**, 848 (2003)].
12. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **80**, 675 (2003) [Astron. Rep. **47**, 621 (2003)].
13. D. Steeghs, *Astrotomography, Indirect Imaging Methods in Observational Astronomy*, Ed. by H. M. J. Boffin, D. Steeghs, and J. Cuypers; Lect. Notes Phys. **573**, 45 (2001).
14. E. T. Harlaftis and T. Marsh, Astron. Astrophys. **308**, 97 (1996).
15. D. V. Bisikalo, A. A. Boyarchuk, P. V. Kaĭgorodov, *et al.*, Astron. Zh. **81**, 494 (2004) [Astron. Rep. **48**, 449 (2004)].
16. D. V. Bisikalo, A. A. Boyarchuk, P. V. Kaigorodov, *et al.*, Astron. Zh. **81**, 648 (2004) [Astron. Rep. **48**, 588 (2004)].
17. O. A. Kuznetsov, D. V. Bisikalo, A. A. Boyarchuk, *et al.*, Astron. Zh. **78**, 997 (2001) [Astron. Rep. **45**, 872 (2001)].
18. D. V. Bisikalo, A. A. Boyarchuk, P. V. Kaigorodov, and O. A. Kuznetsov, Astron. Zh. **80**, 879 (2003) [Astron. Rep. **47**, 809 (2003)].
19. D. O'Donoghue, Mon. Not. R. Astron. Soc. **246**, 29 (1990).
20. N. I. Shakura and R. A. Sunyaev, Astron. Astrophys. **24**, 337 (1973).
21. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **219**, 75 (1986).
22. D. Steeghs, E. T. Harlaftis, and K. Horne, Mon. Not. R. Astron. Soc. **290**, L28 (1997).
23. I. Hachisu, M. Kato, and T. Kato, Astrophys. J. Lett. **606**, L139 (2004).
24. N. I. Balog, A. V. Goncharskiĭ, Z. Yu. Metlitskaya, and A. M. Cherepashchuk, Peremen. Zvezdy **21**, 695 (1982).
25. T. S. Khruzina, Astron. Zh. **78**, 298 (2001) [Astron. Rep. **45**, 255 (2001)].
26. N. A. Katysheva, T. S. Khruzina, S. Yu. Shugarov, *et al.*, *The Physics of Cataclysmic Variables and Related Objects*, Ed. by K. Beuermann, K. Reinsch, and B. Gaensicke; Astron. Soc. Pac. Conf. Ser. **261**, 493 (2002).
27. T. S. Khruzina, Astron. Zh. **75**, 209 (1998) [Astron. Rep. **42**, 180 (1998)].
28. G. Djurasevic, Astrophys. Space Sci. **240**, 317 (1996).
29. I. B. Voloshina and T. S. Khruzina, Astron. Zh. **77**, 109 (2000) [Astron. Rep. **44**, 89 (2000)].
30. S. Wolf, K. H. Mantel, K. Horne, *et al.*, Astron. Astrophys. **273**, 160 (1993).
31. T. R. Marsh and K. Horne, Astrophys. J. **349**, 593 (1990).
32. E. T. Harlaftis, D. Steeghs, K. Horne, *et al.*, Mon. Not. R. Astron. Soc. **306**, 348 (1999).
33. A. Bobinger, H. Barwig, H. Fiedler, *et al.*, Astron. Astrophys. **348**, 145 (1999).
34. R. Baptista, E. T. Harlaftis, and D. Steeghs, Mon. Not. R. Astron. Soc. **314**, 727 (2000).
35. L. Morales-Rueda, T. R. Marsh, and I. Billington, Mon. Not. R. Astron. Soc. **313**, 454 (2000).
36. D. M. Himmelblau, *Applied Nonlinear Programming* (McGraw-Hill, New York, 1972; Mir, Moscow, 1975), p. 163.
37. E. T. Harlaftis, R. Baptista, L. Morales-Rueda, *et al.*, Astron. Astrophys. **417**, 1063 (2004).

Translated by L. Yungel'son

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 801–813. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 900–913.

Dependence of the Absorption-Line Profiles and Radial-Velocity Curve of the Optical Star in an X-ray Binary on the Orbital Inclination and Component-Mass Ratio

M. K. Abubekerov, É. A. Antokhina, and A. M. Cherepashchuk

Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia

Received December 18, 2004; in final form, May 18, 2005

Abstract—Theoretical absorption-line profiles and radial-velocity curves for tidally deformed optical stars in X-ray binary systems are calculated assuming LTE. The variations in the profile shapes and radial-velocity curve of the optical star are analyzed as a function of the orbital inclination of the X-ray binary system. The dependence of the shape of the radial-velocity curve on the orbital inclination i increases with decreasing component-mass ratio $q = m_x/m_v$. The integrated line profiles and radial-velocity curves of the optical star are calculated for the Cyg X-1 binary, which are then used to estimate the orbital inclination and mass of the relativistic object: $i < 43°$ and $m_x = 8.2{-}12.8\,M_\odot$. These estimates are in good agreement with earlier results of fitting the radial-velocity curve of Cyg X-1 using a simpler model ($i < 45°$, $m_x = 9.0{-}13.2\,M_\odot$). © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The optical component in an X-ray binary system is tidally deformed and has a complex temperature distribution on its surface due to the effects of gravitational darkening and X-ray heating. These effects of the interaction between the components give rise to orbital variability of the absorption-line profiles of an optical star. The orbital variability of the absorption profiles leads to a dependence of both the half-amplitude and the shape of the radial-velocity curve on the orbital inclination i and the component-mass ratio $q = m_x/m_v$.

The dependence of the shape of the radial-velocity curve of a tidally deformed optical star on the parameters of a close binary system was first pointed out by Sofia and Wilson [1]. Antokhina and Cherepashchuk [2] and Shabaz [3] proposed a new method for determining the component-mass ratio $q = m_x/m_v$ and orbital inclination i of an X-ray binary based on the orbital variability of the absorption-line profiles in the spectrum of the optical star. Abubekerov *et al.* [4] estimated the orbital inclination of the X-ray binary Cyg X-1 using a high-accuracy observational radial-velocity curve.

Here, we present the results of theoretical modeling of the radial-velocity curve of the optical star in a Roche model assuming LTE for various orbital inclinations, for X-ray systems with low-, moderate-, and high-mass optical stars. We analyze variations in the radial-velocity curve with variations of i for various values of q, together with the corresponding variations in the Hγ absorption-line profile.

2. SYNTHESIS OF THE RADIAL-VELOCITY CURVES FOR THE OPTICAL STAR

The synthesis of the theoretical absorption-line profiles and radial-velocity curves for the optical star in an X-ray binary system was carried out using the algorithm described in detail by Antokhina *et al.* [5, 6]. We will briefly summarize the basis of this method here.

In the Roche model, the X-ray binary system consists of an optical star and a pointlike X-ray source. The star is tidally deformed and has a nonuniform surface-temperature distribution due to the effects of gravitational darkening and heating of the stellar surface by the X-ray emission of the relativistic object.

The surface of the optical star was divided into ~2600 area elements, for each of which we calculated the emergent local radiation assuming LTE. Each area element corresponds to a local temperature $T_{\rm loc}$, local gravitational acceleration $g_{\rm loc}$, and local value of the parameter $k_x^{\rm loc}$, which is equal to the ratio of the incident X-ray flux and the outgoing radiation flux without allowance for external irradiation of the atmosphere. Using these parameter values at a given point of the surface, a model for the atmosphere is calculated by solving the equation of line radiative transfer in the presence of incident external X-ray radiation. In this way, we can compute the intensity of the outgoing radiation in the line and continuum for each local area element. At different phases of the orbital period, the contributions of the areas to the total radiation are summed taking into account Doppler

1063-7729/05/4910-0801$26.00

Table 1. Numerical values of parameters used to synthesize the radial-velocity curves of the optical component in the Roche model

P, day	5.0	Period
m_x, $M_\odot$	10	Mass of the compact object
m_v, $M_\odot$	*var**	Mass of the optical star
e	0.0	Eccentricity
i, deg	30, 60, 90	Orbital inclination
μ	1.0	Roche lobe filling coefficient for the optical star
f	1.0	Asynchronicity coefficient for the rotation of the optical star
$T_{\rm eff}$, K	*var**	Effective temperature of the optical star
β	0.25**	Gravitational-darkening coefficient
k_x	0.1	Ratio of the X-ray luminosity of the relativistic component to the bolometric luminosity of the optical component L_x/L_v
A	0.5	Reprocessing coefficient for the X-ray radiation
u	0.3	Limb-darkening coefficient

* Parameter of the X-ray binary was varied during the model computations.
** When the mass of the optical star was $m_v = 1\,M_\odot$, the gravitational-darkening coefficient was taken to be $\beta = 0.08$.

effects and the conditions for visibility of the areas by an observer. This yields the total radiation flux from the star in the direction toward the observer in the continuum and a rotationally broadened spectral-line profile that can be used to derive the radial velocity of the star.

Since the observed radial-velocity curves for OB stars have been derived primarily from hydrogen Balmer absorption lines, we synthesized theoretical radial-velocity curves for the Hγ line of the optical star. The radial velocity at a given orbital phase was calculated using the mean wavelength at the 1/3, 1/2, and 2/3 levels of the maximum depth of the integrated absorption-line profile.

In addition, in order to estimate the uncertainties in the modeling, we carried out computations for the same situations using our previous algorithm for synthesizing theoretical radial-velocity curves, used earlier to analyze the radial-velocity curves of OB supergiants in X-ray binaries with neutron stars [7] and in the Cyg X-1 system [4]. This algorithm was proposed by Antokhina and Cherepashchuk [8] in 1994. The main difference from our more modern algorithm [5, 6] is that the local profile of an area element is found, not by constructing a model atmosphere and calculating the intensity of the outgoing radiation in the line and continuum (taking into account reprocessing of the external X-ray radiation), but instead using computed and tabulated Kurucz Balmer absorption-line profiles for various effective temperatures $T_{\rm eff}$ and gravitational accelerations g. In addition, the effect of heating of the stellar atmosphere by X-ray radiation from its companion was taken into account only in a simple way, by adding the outgoing and incident flux units without taking into account radiative transfer in the stellar atmosphere.

In addition to its simplified treatment of the reflection effect, this method for calculating the hydrogen absorption-line profiles is not entirely correct, since the tables of Kurucz [9] present the theoretical line profiles in relative flux units and not intensities. However, since we are using the theoretical line profiles to derive radial velocities, and not for comparisons with observed spectral lines, we consider this approximation to be acceptable. In addition, calculating the theoretical radial-velocity curves using the algorithm described based on the tables of Kurucz requires comparatively little computer time. The use of the more modern algorithm [5, 6], which calculates a model atmosphere for each local area element, requires appreciably more computer time and became possible only relatively recently with the appearance of computers with processing rates of 1 GHz or higher.

One drawback of the new algorithm is the absence of a contribution to the integrated absorption-line profile of the optical star from the local profiles of areas near the limb of the stellar disk (the requirement of

Table 2. Mass and corresponding effective temperature of the optical star

$m_v, M_\odot$	$T^*_{\rm eff}$, K	$T^{**}_{\rm eff}$, K	$T^{***}_{\rm eff}$, K	Range of atmospheric parameter values on the stellar surface in the Roche model		S, %
				$\log g$	$T_{\rm eff}$, K	
1	–	5500	5500	2.02–3.15	2800–5930	89
5	14 000	12 000	12 000	2.02–3.16	6570–12 820	93
10	17 000	15 000	16 000	2.13–3.22	9190–17 060	94
20	26 000	23 000	23 000	2.23–3.25	14 020–24 540	94
30	29 000	28 000	29 000	2.28–3.26	18 390–31 010	96

Note: $T^*_{\rm eff}$ and $T^{**}_{\rm eff}$ are the temperatures of the optical star according to the mass–luminosity relations of [11] and [12]; $T^{***}_{\rm eff}$ is the temperature of the optical star used to compute the radial-velocity curves in the current study; and S the fraction of the area of the stellar surface where the temperature differs from the mean effective temperature $T^{***}_{\rm eff}$ by no more than 10%.

the boundary conditions [5, 6]). At the same time, for real optical stars, the local profiles of areas near the disk limb are nonzero and will make a significant contribution to an integrated absorption-line profile. Note that the old algorithm was free of this effect of "zeroing" the contributions of the local profiles for areas near the disk limb.

We used both these algorithms to synthesize the theoretical radial-velocity curves for all the model problems considered. The results of these computations using these two methods are compared below.

Table 1 presents the parameters of the X-ray binary for which the modeling was carried out. The numerical values of the parameters of the modeled close binary were adopted based on the catalog [10], as being the most characteristic values.

We synthesized radial-velocity curves for optical stars with masses $m_v = 1,\ 5,\ 10,\ 20,\ 30\,M_\odot$ (the remaining parameters of the modeled binary system are presented in Table 1). In order to investigate the dependence of the shape of the radial-velocity curve on the orbital inclination i, we synthesized radial-velocity curves for $i = 30°$, 60°, and 90°.

When modeling the radial-velocity curves of an optical star with mass $m_v = 30\,M_\odot$ using the tables of Kurucz [9], the local gravitational acceleration $g_{\rm loc}$ and local temperature $T_{\rm loc}$ at the surface of the optical star fell outside the range of tabulated values, so that there were no tabulated profiles for some of the local areas. The number of such areas was modest (about 10–20% of the total) and they were all located on the "nose" of the filled Roche lobe of the optical component. In this case, the profiles of all the areas were taken to be the same. We used the profile of the Hγ line for the mean effective temperature and the mean gravitational acceleration at the stellar surface as the local profile. The accuracy of this approximation was tested in test computation 1 presented below.

We should also qualify the situation with regard to the mean effective temperature of the optical star. The mean temperature of the Roche lobe filling star is given by the expression $T_{\rm eff} = \int T_{\rm loc} dS / \int dS$, where the integration is carried out over the entire surface of the tidally deformed star. Table 2 presents estimates of the effective temperature based on the mass–luminosity relations of Herrero [11] and Straizhis [12]. We can see that the effective temperatures of the star obtained in different ways are appreciably different. Therefore, we carried out a test computation (test computation 2) in order to quantitatively estimate the influence of the effective temperature of the optical star on the theoretical radial-velocity curve. We took the mean effective temperature of the tidally deformed star to be the mean effective temperature of a spherical star with the same volume. The results of the test computations are presented below.

Table 2 also presents the range of effective temperatures and gravitational accelerations at the surface of the optical star in the Roche model. Note that the number of local area elements whose temperature $T_{\rm loc}$ differs from the adopted effective temperature $T^{***}_{\rm eff}$ by more than 10% is very small.

Test Computation 1: Influence of the Local-Profile Approximation on the Theoretical Radial-Velocity Curve

As we indicated above, the test computations check the difference in the radial velocities calculated using the two methods and the tables of Kurucz [9]. We synthesized the radial-velocity curve for an optical

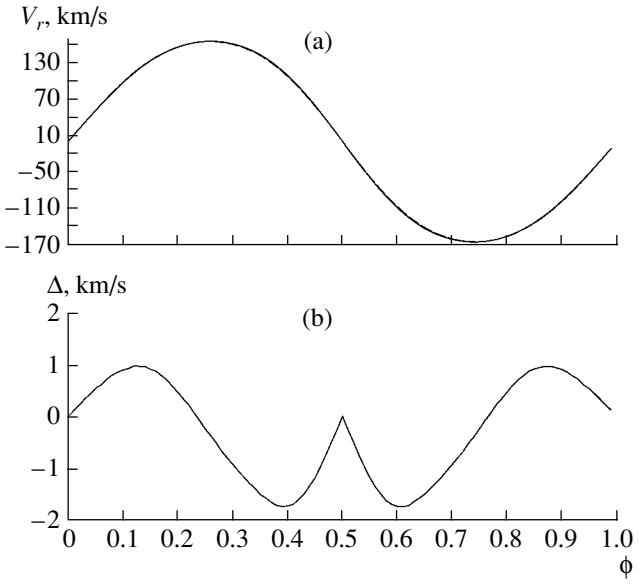

Fig. 1. Test computation 1. (a) Model radial-velocity curve for an optical star with $m_v = 10\,M_\odot$ and $T_{\rm eff} = 15\,000$ K for $i = 90^\circ$ (solid) together with the same curve obtained assuming that the shape of the local profile is constant (dashed). The two radial-velocity curves virtually coincide on the scale shown. These curves were calculated using the tabulated Hγ line profiles (in flux units) of Kurucz. (b) Difference between the absolute values of the radial velocities. See text for more detail.

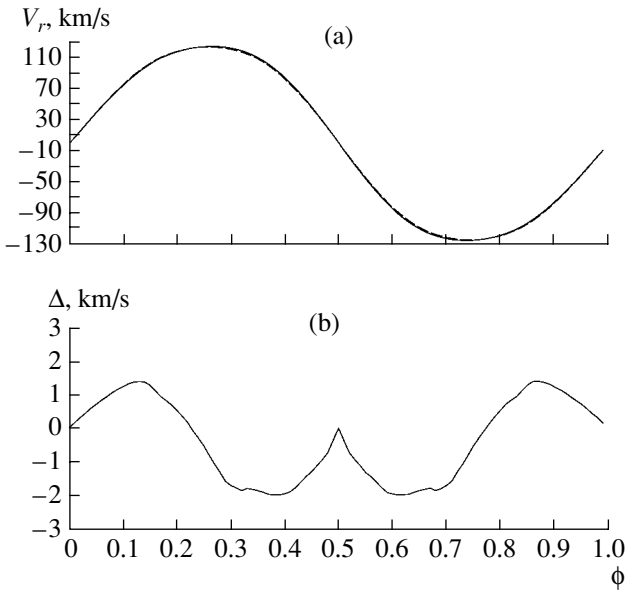

Fig. 2. Same as Fig. 1 for the parameters of the optical star $m_v = 20\,M_\odot$ and $T_{\rm eff} = 23\,000$ K.

star with mass $m_v = 10\,M_\odot$ and mean effective surface temperature $T_{\rm eff} = 15\,000$ K (the remaining binary parameters can be found in Table 1); each local area is specified in the LTE approximation in accordance with its local Hγ absorption profile from the tables of Kurucz (solid curve in Fig. 1a). We then constructed the radial-velocity curve for the optical star specifying the shapes of the local profiles for all the area elements to be the same, but taking into account normalization to the continuum over the stellar surface when summing the area profiles. For the constant profile shape, we used the Hγ absorption profile for the mean effective temperature and gravitational acceleration of the optical star. In the case of a star with $m_v = 10\,M_\odot$, we used the tabulated profile of Kurucz [9] corresponding to $T_{\rm loc} = 15\,000$ K and $\log g_{\rm loc} = 3.2$. The resulting radial-velocity curve is shown by the dotted curve in Fig. 1a. Figure 1b shows the difference between the absolute values of the radial velocities obtained using the "intrinsic" local profiles and assuming identical local profiles for each area element. We can see that the difference between the curves does not exceed 1.7 km/s, or $\sim$1% of the half-amplitude of the radial-velocity curve. An analogous computation was carried out for a close X-ray binary with an optical star with mass $m_v = 20\,M_\odot$. The resulting radial-velocity curves are shown in Fig. 2a. In this case, the discrepancy between the radial-velocity curves does not exceed 2 km/s, or 1.6% of the radial-velocity half-amplitude (Fig. 2b).

Test Computation 2: Influence of the Effective Temperature of the Optical Star on the Theoretical Radial-Velocity Curve

As we noted above, the effective temperature of the optical star is usually not known exactly. We can see from Table 2 that the observed effective temperatures derived from the mass–luminosity relations of [11, 12] are somewhat different. Therefore, we carried out a test computation to estimate the influence of the effective temperature of the optical star on the theoretical radial-velocity curve. We synthesized a radial-velocity curve for a close binary with an optical star with mass $m_v = 10\,M_\odot$ and orbital inclination $i = 90^\circ$ for $T_{\rm eff} = 10\,000$ K and $T_{\rm eff} = 17\,000$ K (having especially chosen a wide range of variation for the effective temperature). The resulting radial-velocity curves are shown in Fig. 3a (see Table 1 for the remaining binary parameters). Since the discrepancy between the curves is insignificant, Fig. 3b presents the difference between their absolute values, $\Delta = |V_r(T_{\rm eff} = 17\,000\ \ {\rm K})| - |V_r(T_{\rm eff} = 10\,000\ {\rm K})|$, where $|V_r(T_{\rm eff} = 17\,000\ {\rm K})|$ and $|V_r(T_{\rm eff} = 10\,000\ {\rm K})|$ are the absolute values of the radial velocity of the optical star

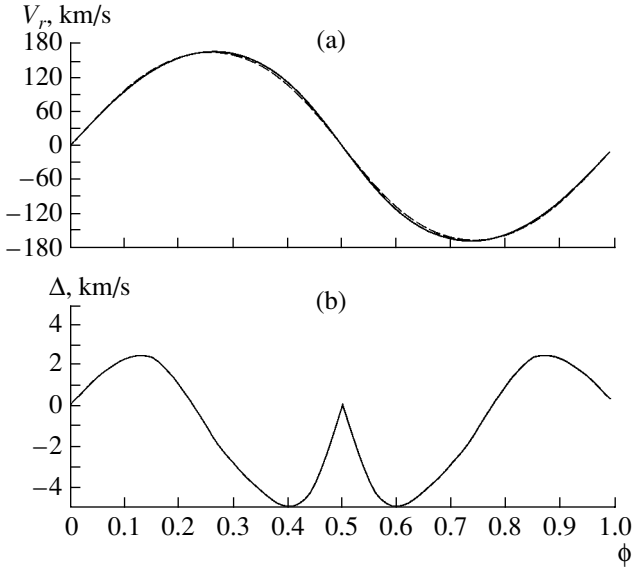

Fig. 3. Test computation 2. (a) Model radial-velocity curves of an optical star with $m_v = 10\,M_\odot$, $i = 90°$, and $T_{\rm eff} = 10\,000$ K (solid) and $T_{\rm eff} = 17\,000$ K (dashed). The curves were calculated using the tabulated Hγ profiles (in flux units) of Kurucz. (b) Difference between the absolute values of the radial velocities $\Delta = |V_r(T_{\rm eff} = 17\,000\ {\rm K})| - |V_r(T_{\rm eff} = 10\,000\ {\rm K})|$. See text for more detail.

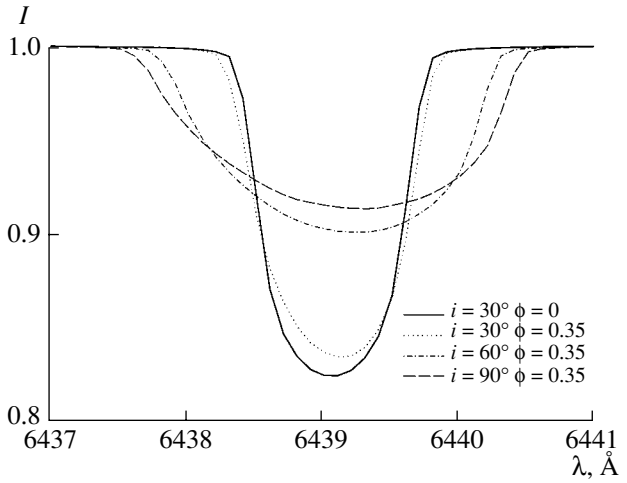

Fig. 4. Theoretical CaI 6439 Å absorption profiles without allowance for an instrumental profile (calculated with the new algorithm [5, 6]) at orbital phase $\phi = 0.0$ for orbital inclination $i = 30°$ (solid) and at orbital phase $\phi = 0.35$ for $i = 30°$ (dotted), 60° (dash-dotted), and 90° (dashed). The model profiles were obtained assuming LTE and a mass and effective temperature for the optical star $m_v = 1\,M_\odot$ and $T_{\rm eff} = 5500$ K (see Table 1 for the remaining parameters). The CaI 6439 Å absorption profiles for phase 0.35 have been corrected for the Doppler shifts due to the orbital motion.

in the Roche model for mean effective surface temperatures of $T_{\rm eff} = 17\,000$ K and $T_{\rm eff} = 10\,000$ K. We can see from Fig. 3b that the maximum discrepancy between these two values occurs at orbital phases 0.35–0.45, and reaches 5 km/s, or 2.6% of the radial-velocity half-amplitude.

We can see from the test computation that the uncertainty in the mean effective temperature of the optical star ($\Delta T_{\rm eff} \simeq 5000$–$7000$ K) has an appreciable effect on the shape of the radial-velocity curve. The variations in the radial-velocity curve are different at different orbital phases (Fig. 3b). These variations are maximum at phases 0.35–0.45, where they reach ∼3% of the radial-velocity half-amplitude. As we noted above, the uncertainty in the effective temperature was artificially increased in the test computation. According to Table 2, the maximum uncertainty in the effective temperature does not exceed 3000 K, so that the corresponding variations in the radial-velocity curve should not exceed ∼1%.

Our computations indicate that the mean effective temperature of the optical star, $T_{\rm eff}$, should be known as accurately as possible when using radial-velocity curves to determine the parameters of a binary system.

3. DEPENDENCE OF THE SHAPE OF THE RADIAL-VELOCITY CURVE ON THE COMPONENT-MASS RATIO q AND ORBITAL INCLINATION i

We also calculated radial-velocity curves with and without allowing for the influence of the instrumental function on the theoretical integrated profile.

Calculated Radial-Velocity Curves without Account of the Instrumental Function on the Theoretical Absorption Profile

We synthesized radial-velocity curves in the Roche model assuming LTE for optical stars with masses $m_v = 1, 5, 10, 20$, and $30\,M_\odot$ (see Table 1 for the remaining parameters of the X-ray binary). For the optical star with mass $m_v = 1\,M_\odot$, the synthesis was carried out for the CaI 6439 Å absorption line. The radial-velocity curve syntheses for the stars with masses $m_v = 5$, 10, 20, and $30\,M_\odot$ were performed for the Hγ absorption line using the two methods described above (calculating the intensity of the local profile of each area element based on a constructed model atmosphere [5, 6] and based on the tabulated line profiles in flux units of Kurucz [8, 13]). The theoretical integrated CaI 6439 Å and Hγ absorption profiles are presented in Figs. 4 and 5. For each value of m_v, we synthesized radial-velocity curves for

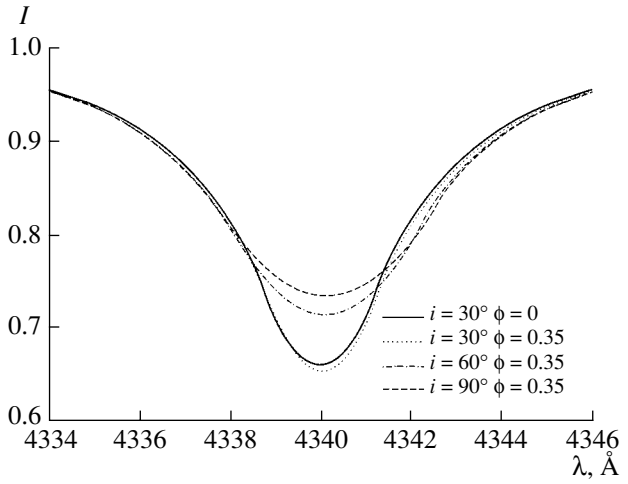

Fig. 5. Same as Fig. 4 for the Hγ absorption line with $m_v = 20\,M_\odot$ and $T_{\rm eff} = 23\,000$ K.

orbital inclinations $i = 30°$, $i = 60°$, and $i = 90°$. The resulting theoretical radial-velocity curves for optical stars with masses $m_v = 1\,M_\odot$, $20\,M_\odot$, and $30\,M_\odot$ are presented in Figs. 6a, 7a, and 8a.

The half-amplitude of the radial-velocity curve K_v increases with increasing orbital inclination i (Figs. 6a, 7a, 8a). Therefore, each theoretical radial-velocity curve was normalized to its value of K_v, which was taken to be the maximum radial velocity in the orbital-phase interval 0.0–0.5. The normalized radial-velocity curves for optical stars with masses $m_v = 1$, 20, $30\,M_\odot$ are presented in Figs. 6b, 7b, 8b. Since the effect of the differences in the shapes of the curves is small (relative to the half-amplitude K_v), we calculated the differences of the absolute values of the normalized radial velocities $\Delta = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 30°)|$ and $\Delta = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 60°)|$, where $|V_{\rm norm}(i = 90°)|$, $|V_{\rm norm}(i = 60°)|$, and $|V_{\rm norm}(i = 30°)|$ are the absolute values of the normalized radial velocities for orbital inclinations of $i = 90°$, $i = 60°$, and $i = 30°$. The maximum differences between the normalized radial-velocity curves obtained using the two methods for optical stars with masses $m_v = 1$, 20, $30\,M_\odot$ are reached at orbital phases 0.35–0.45 and 0.55–0.65 (Tables 3 and 4). We denote Δ_I and Δ_F to be the maximum variations in the shape of the radial-velocity curve for variations in the orbital inclination obtained using the new algorithm of [5, 6] and the theoretical line profiles of Kurucz [8, 13], respectively.

The difference in the absolute values of the normalized radial velocities for optical stars with masses $m_v = 1$, 20, $30\,M_\odot$ are presented in Figs. 6c, 7c, 8c. We can see from these figures that the maximum difference between the radial-velocity curves obtained

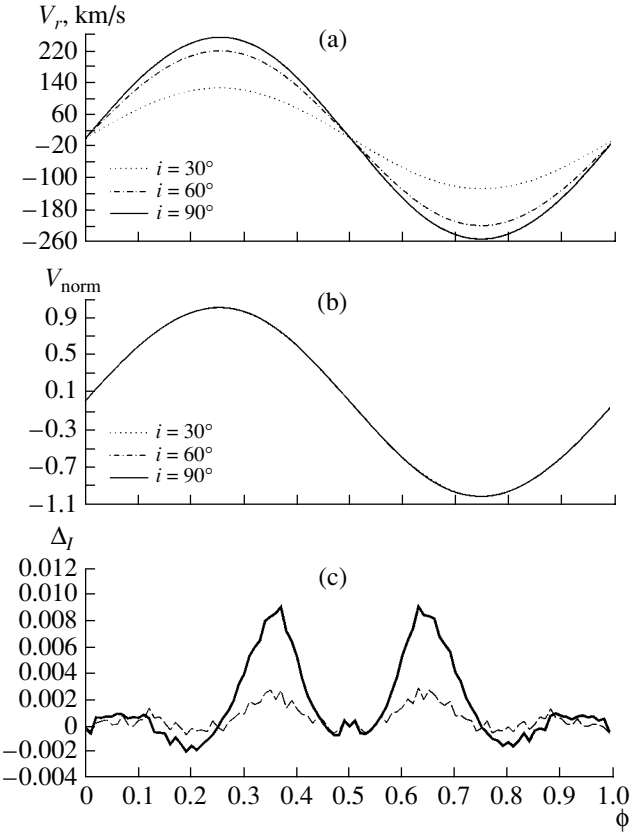

Fig. 6. (a) Theoretical radial-velocity curve of an optical star with $m_v = 1\,M_\odot$, $T_{\rm eff} = 5500$ K calculated without allowance for the effect of the instrumental function on the integrated CaI 6439 Å absorption profile for orbital inclinations $i = 30°$ (dotted), 60° (dash-dotted), and 90° (solid) (see Table 1 for remaining parameters). The radial-velocity curves were calculated using the new algorithm [5, 6]. (b) The same curves normalized to their half-amplitudes for $i = 30°$ (dotted), 60° (dash-dotted), and 90° (solid). The curves nearly coincide on this scale. (c) Difference between the absolute values of the normalized radial velocities $\Delta_I = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 60°)|$ (dashed) and $\Delta_I = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 30°)|$ (solid). The quantity Δ_I is given in units of the half-amplitude of the radial-velocity curve K_v (for more detail, see text).

for different values of i occurs at orbital phases 0.35–0.45 and 0.55–0.65.

It follows from the computations that the variations in the shape of the radial-velocity curve depend on both the orbital inclination i and the component-mass ratio $q = m_x/m_v$. For example, the maximum variations in this curve (in units of the radial-velocity half-amplitude) when the inclination is changed from $i = 30°$ to $i = 90°$ for optical stars with masses $m_v =$ 1, 5, 10, 20, and $30\,M_\odot$ are $\Delta_F = 0.9$, 3.2, 4.6, 5.7, and 8%, respectively.

Thus, the variations in the shape of the radial-

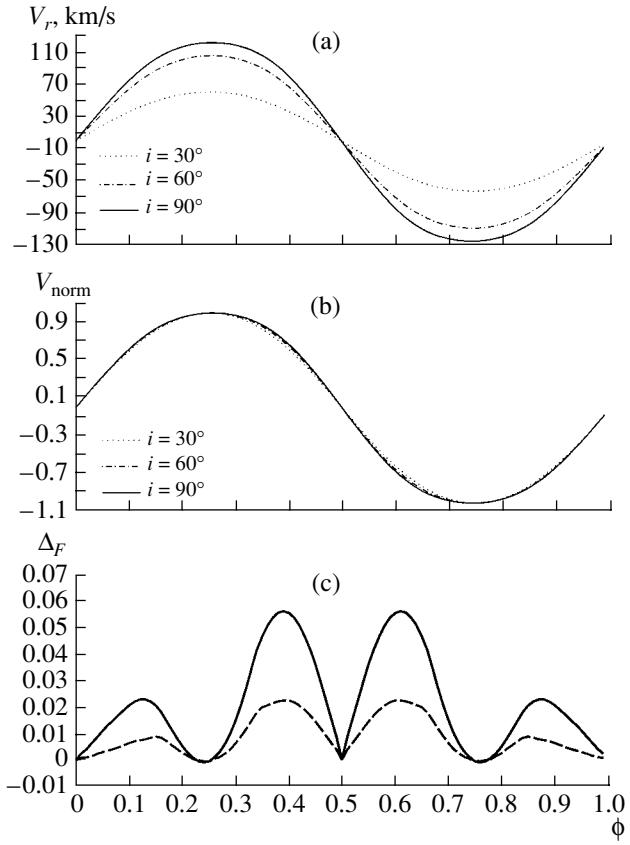

Fig. 7. (a) Theoretical radial-velocity curve of an optical star with $m_v = 20\,M_\odot$, $T_{\rm eff} = 23\,000$ K calculated without allowance for the effect of the instrumental function on the integrated Hγ absorption profile for orbital inclinations $i = 30°$ (dotted), 60° (dash-dotted), and 90° (solid) (see Table 1 for remaining parameters). The radial-velocity curves were calculated using the old algorithm [8, 13] (using the tabulated Hγ absorption profiles of Kurucz in flux units). (b) The same curves normalized to their half-amplitudes for $i = 30°$ (dotted), 60° (dash-dotted), and 90° (solid). (c) Difference between the absolute values of the normalized radial velocities $\Delta_F = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 60°)|$ (dashed) and $\Delta_F = |V_{\rm norm}(i = 90°)| - |V_{\rm norm}(i = 30°)|$ (solid). The quantity Δ_F is given in units of the half-amplitude of the radial-velocity curve K_v (for more detail see text).

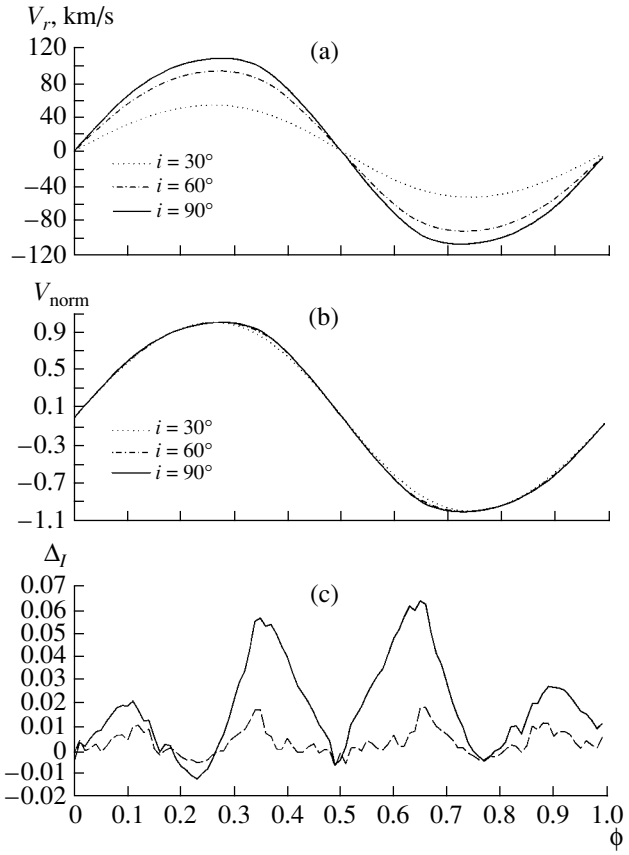

Fig. 8. Same as Fig. 6 for the Hγ absorption line and with $m_v = 30\,M_\odot$ and $T_{\rm eff} = 29\,000$ K.

velocity curve with changes in the orbital inclination are most clearly expressed in systems with low component-mass ratios $q = m_x/m_v$. This is due to the fact that, when $q < 1$, the center of mass of the binary system lies inside the body of the optical star, so that, during the orbital motion, the part of the stellar surface lying inside the inner Lagrangian point moves in the same direction as the relativistic object. This leads to strong distortion of the integrated absorption-line profile that depends on both i and q.

Computation of Radial-Velocity Curves Including the Effect of the Instrumental Function on the Theoretical Absorption Profiles

As in the previous case, the radial-velocity curve synthesis was carried out for optical stars with masses $m_v = 1$, 5, 10, 20, and $30\,M_\odot$ (see Table 1 for the remaining binary parameters). We used a Gaussian profile for the instrumental function of the spectrograph. The CaI 6439 Å line was convolved with an instrumental function with a full width at half maximum intensity FWHM = 1 Å. When synthesizing the curves for stars with masses $m_v =$ 5, 10, 20, and $30\,M_\odot$, the theoretical Hγ profile was convolved with an instrumental function with FWHM = 7 Å. The convolved theoretical CaI 6439 Å and Hγ absorption profiles are presented in Figs. 9 and 10, respectively.

We investigated the variations of the radial-velocity curves with variations of the orbital inclination analogous to those described above. As earlier, the maximum changes in the curves occurred at phases 0.35–0.45. The variations in the radial-velocity curve obtained when the inclination was changed from $i = 60°$ to $i = 90°$ are given in Table 5.

Table 3. Maximum variation of the shape of the radial-velocity curve in the Roche model for variations of the orbital inclination from 60° to 90°

$m_v, M_\odot$	$\Delta_I, \%$	$\Delta_F, \%$
1	0.3	–
5	0.4	1.1
10	0.6	1.5
20	1.1	2.3
30	1.9	2.9

Note: The quantities Δ_I and Δ_F are expressed in units of the half-amplitude K_v. See text for more detail.

Table 4. Maximum variation of the shape of the radial-velocity curve in the Roche model for variations of the orbital inclination from 30° to 90°

$m_v, M_\odot$	$\Delta_I, \%$	$\Delta_F, \%$
1	0.9	–
5	1.5	3.2
10	2.5	4.6
20	4.2	5.7
30	6.5	8.0

Note: The quantities Δ_I and Δ_F are expressed in units of the half-amplitude K_v. See text for more detail.

The analogous results for the case when the inclination is changed from $i = 30°$ to $i = 90°$ are given in Table 6. A comparison of Tables 3, 4 and 5, 6 shows that the results obtained with and without the effect of the instrumental profile coincide within reasonable errors for the computed integrated profiles.

Thus, the variations in the shape of the radial-velocity curve with orbital inclination are not "washed out" by the instrumental function of the spectrograph, leaving intact the possibility of estimating the orbital inclination of a binary system using accurate observations of the radial-velocity curve. Estimating the orbital inclination of a system with a given mass for its optical star reduces to a required accuracy for the radial-velocity observations. For example, for binary systems with optical-star masses $m_v = 20–30\,M_\odot$ (close to filling their Roche lobes), the accuracy in the radial velocities must be better than 7–8% of the half-amplitude of the radial-velocity curve, K_v. For systems with optical-star masses $m_v = 1–10\,M_\odot$ (close to filling their Roche lobes), this accuracy must be better than 1–4% of K_v. The orbital inclination of the Cyg X-1 system estimated using an accurate radial-velocity curve is presented in [4]; the accuracy in the observed radial velocities σ_{V_r} was ~3% of the half-amplitude K_v. For an optical-star mass of $m_v \simeq 20\,M_\odot$, this made it possible to place a limit on the orbital inclination of $i < 45°$ based purely on the observed radial-velocity curve.

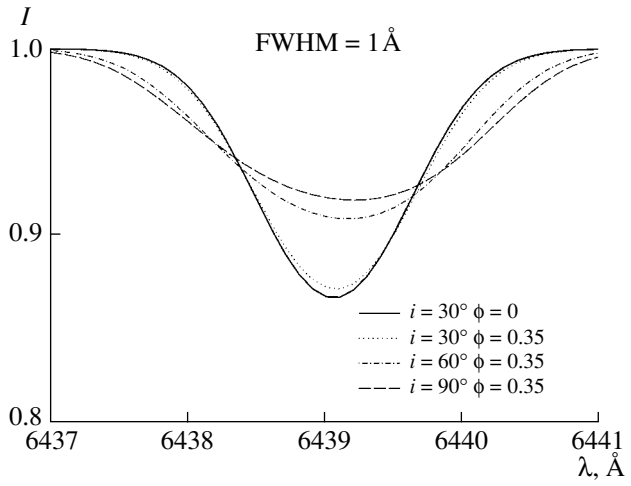

Fig. 9. Same as Fig. 4 but including the effect of an instrumental profile with FWHM = 1 Å.

4. FITTING THE OBSERVED RADIAL-VELOCITY CURVE OF THE Cyg X-1 SYSTEM

The fitting of the mean radial-velocity curve for Cyg X-1 presented in [4] was carried out with the earlier algorithm, which used the tabulated Hγ profiles of Kurucz [9] in fluxes for the local profiles of the area elements. However, as we can see from Tables 3–6, the effects of varying the shapes of the radial-velocity curves calculated using the old and new algorithms are somewhat different. For this reason, we decided to fit the accurate mean radial-velocity curve considered in [4] using the new algorithm, in which the integrated Hγ absorption profiles are calculated in intensity units based on model atmospheres for the local area elements [5, 6]. The calculations based on the old algorithm yielded a mass for the compact object in the Cyg X-1 binary of $m_x = 9.0–13.2\,M_\odot$ [4]. No model for the binary system with an orbital inclination $i < 45°$ is consistent with the accurate observational radial-velocity curve [4].

We used the Fisher statistical criterion to test the adequacy of the model description of the observational data. The method used is described in detail in [4], and we do not present it here. As in our earlier work [4], the test of the model description of the observational data was made for the $\alpha = 5\%$ significance level.

Table 5. Maximum variation of the shape of the radial-velocity curve in the Roche model for variations of the orbital inclination from 60° to 90°, including the effect of the instrumental function on the theoretical integrated absorption profile

$m_v, M_\odot$	Δ_I, %	Δ_F, %
1	0.3	–
5	0.4	1.1
10	0.6	1.4
20	1.1	1.7
30	2.0	3.0

Note: The quantities Δ_I and Δ_F are expressed in units of the half-amplitude K_v. See text for more detail.

Table 6. Maximum variation of the shape of the radial-velocity curve in the Roche model for variations of the orbital inclination from 30° to 90°, including the effect of the instrumental function on the theoretical integrated absorption profile

$m_v, M_\odot$	Δ_I, %	Δ_F, %
1	0.9	–
5	1.6	3.3
10	2.4	4.2
20	4.2	5.5
30	6.8	8.7

Note: The quantities Δ_I and Δ_F are expressed in units of the half-amplitude K_v. See text for more detail.

Fits of the radial-velocity curve using the new algorithm were obtained both with and without including the effect of the instrumental function on the model integrated Hγ absorption profile of the optical star. The width of the instrumental function was taken to be FWHM = 7 Å. Fitting was carried out for masses of the optical star $m_v = 20$, 30, 40, 50, and $60\,M_\odot$; the remaining model parameters are given in Table 2 in [4]. This analysis yielded dependences between the masses of the optical star and of the relativistic object for orbital inclinations $i = 30°$, 35°, 40°, 45°, 55°, and 65°, presented in Fig. 11.

As before, the results for orbital inclinations $i < 45°$ are rejected at the 5% significance level. The mass estimates obtained are presented in Tables 7 and 8.

If the mass of the optical star is $m_v = 22\,M_\odot$ [4, 11], the orbital inclination does not exceed $i = 43°$ (Fig. 11). Since analysis of the photometric light curve places a lower limit on the orbital inclination of $i > 31°$ [14], we can conclude that the mass of the relativistic object in the Cyg X-1 binary (for a mass of the optical star $m_v = 22\,M_\odot$) lies in the ranges $m_x = 8.4$–$12.8\,M_\odot$ (Fig. 11a) and $m_x = 8.2$–$12.6\,M_\odot$ (Fig. 11b). Recall that our earlier results for Cyg X-1 [4] indicated an orbital inclination $i < 45°$ and a mass of the relativistic object $m_x = 9.0$–$13.2\,M_\odot$ (for a mass of the optical star $m_v = 22\,M_\odot$). Thus, the results obtained for the new (more correct) [5, 6] and old [8, 13] algorithms are in good agreement.

Based on the example of the Cyg X-1 system, we can see that the accuracy of the mean observed radial-velocity curve, 3% (of a half-amplitude of ∼75 km/s), was sufficient to place constraints on the orbital inclination of the binary system and estimate the mass of the black hole based on a single radial-velocity curve.

5. DISCUSSION

The model computations show that the variations in the shape of the radial-velocity curve of the optical star with variations in the orbital inclination are more clearly expressed in binaries with massive optical components (or low component-mass ratios $q = m_x/m_v$). In addition, these variations are qualitatively similar for optical stars with low, moderate, and high masses—the largest variations (in units of the half-amplitude K_v) occur at orbital phases 0.35–0.45 and 0.55–0.65.

The variations in the shape of the radial-velocity curve are related to variations in the absorption-line profiles used to determine the radial velocities. Recall that the radial velocities were determined for the mean wavelength found for three levels: 1/3, 1/2, and 2/3 of the maximum depth of the absorption line. The

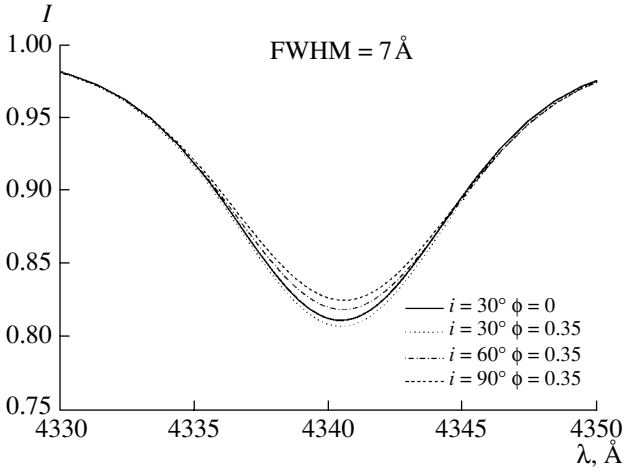

Fig. 10. Same as Fig. 5 but including the effect of an instrumental profile with FWHM = 7 Å.

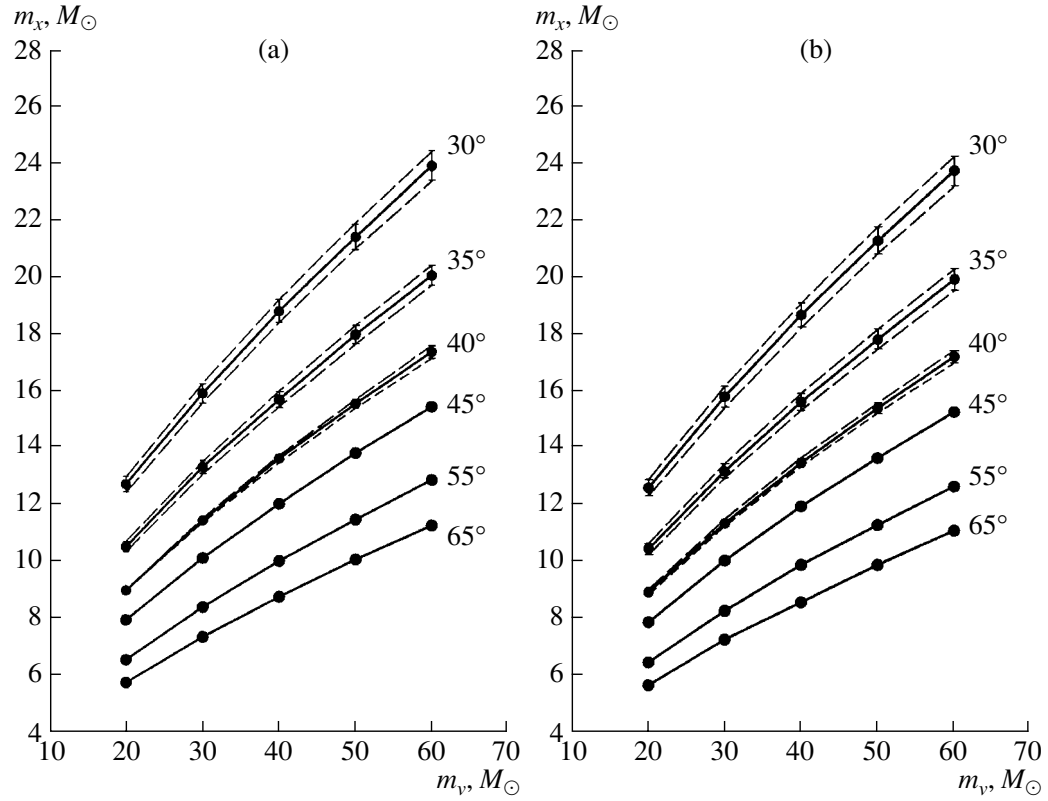

Fig. 11. (a) Dependence of the mass of the compact object in the X-ray binary Cyg X-1 on the mass of the optical star for orbital inclinations $i = 30°$, $35°$, $40°$, $45°$, $55°$, and $65°$, obtained in the Roche model fitting the mean high-accuracy radial-velocity curve from [4]. The orbital inclination is indicated next to the corresponding curve. (b) Same as (a) but including the effect of an instrumental profile with FWHM = 7 Å on the calculated integrated Hγ profile of the optical star. The new algorithm of [5, 6] was used when calculating the integrated Hγ profiles and radial-velocity curves.

shape of an absorption line associated with the optical star in a close binary undergoes variations due to both its orbital motion and changes in the line of sight from the observer toward the system. The dependence of the shape of such an absorption profile due to the star's orbital motion is discussed in [1–3] and is associated with variations in the shape of the projection of the optical star onto the plane of the sky. Let us now consider how variations in the binary orbital inclination i affect the shape of the absorption-line profiles and radial-velocity curve.

An optical star in a binary that is close to filling its Roche lobe is tidally deformed and has complex distributions of the velocity, temperature, and gravitational acceleration over its surface. Therefore, regions of the optical component with qualitatively different parameters are projected onto the plane of the sky of the observer when the orbital inclination i is varied (Fig. 12). This leads to variations in the absorption profiles, which, in turn, lead to variations in the radial-velocity curve. Figures 4 and 5 present the integrated absorption profiles for the CaI 6439 Å and Hγ lines obtained without including the effect of the instrumental profile for orbital inclinations $i = 30°$, $60°$, and $90°$. It is clear that the profiles obtained for the different orbital inclinations are different. The profiles are appreciably asymmetric, and this asymmetry increases with increasing values of i.

The dependence of the absorption profiles and radial-velocity curve on i grows as the component-mass ratio q decreases. Returning to the computational results, we see that, when $q = 1–10$, the variations in the radial-velocity curve are $\sim 1–4\%$ (Fig. 6), while, when $q = 0.3–0.5$, these variations increase to $\sim 4–8\%$ (Figs. 7 and 8). It is clear that the magnitude of the dependence of the radial-velocity curve on i grows with decreasing $q = m_x/m_v$. This is due to the spatial disposition of the center of mass of the binary relative to the body of the optical star. As q decreases, the barycenter of the binary system shifts toward the optical component, with the center of mass lying inside the body of the optical star when $q < 1$. In this case ($q < 1$), the nose of the optical star moves in the same direction along the orbit as the

Table 7. Mass of the relativistic component as a function of the mass of the optical star for the Roche model and orbital inclinations $i = 30°$, $35°$, $40°$, $45°$, $55°$, and $65°$ based on the model integrated Hγ profiles, without including the instrumental function

$m_v, M_\odot$	$m_x, M_\odot$					
	$i = 30°$	$i = 35°$	$i = 40°$	$i = 45°$	$i = 55°$	$i = 65°$
20	$12.75^{+0.25}_{-0.30}$	$10.55^{+0.19}_{-0.17}$	$9.03^{+0.02}_{-0.02}$	8.0	6.60	5.80
30	$15.97^{+0.36}_{-0.32}$	$13.35^{+0.20}_{-0.25}$	$11.47^{+0.08}_{-0.07}$	10.15	8.45	7.40
40	$18.85^{+0.40}_{-0.40}$	$15.75^{+0.30}_{-0.25}$	$13.65^{+0.10}_{-0.15}$	12.05	10.05	8.80
50	$21.45^{+0.50}_{-0.40}$	$18.03^{+0.32}_{-0.33}$	$15.60^{+0.15}_{-0.15}$	13.85	11.50	10.10
60	$23.95^{+0.50}_{-0.55}$	$20.10^{+0.35}_{-0.35}$	$17.43^{+0.22}_{-0.23}$	15.50	12.90	11.30

Note: The confidence intervals for m_x for $i = 45°$, $55°$, and $65°$ are not indicated, since the model of the binary system is rejected at the $\alpha = 5\%$ significance level in these cases.

Table 8. Mass of the relativistic component as a function of the mass of the optical star for the Roche model and orbital inclinations $i = 30°$, $35°$, $40°$, $45°$, $55°$, and $65°$ based on the model integrated Hγ profiles, including an instrumental function with FWHM = 7 Å

$m_v, M_\odot$	$m_x, M_\odot$					
	$i = 30°$	$i = 35°$	$i = 40°$	$i = 45°$	$i = 55°$	$i = 65°$
20	$12.60^{+0.30}_{-0.28}$	$10.45^{+0.20}_{-0.20}$	$8.95^{+0.10}_{-0.10}$	7.9	6.50	5.70
30	$15.83^{+0.37}_{-0.38}$	$13.18^{+0.27}_{-0.23}$	$11.35^{+0.15}_{-0.10}$	10.05	8.30	7.30
40	$18.70^{+0.40}_{-0.45}$	$15.65^{+0.30}_{-0.30}$	$13.48^{+0.17}_{-0.13}$	11.95	9.90	8.60
50	$21.30^{+0.50}_{-0.45}$	$17.85^{+0.35}_{-0.35}$	$15.43^{+0.17}_{-0.18}$	13.65	11.30	9.90
60	$23.75^{+0.50}_{-0.55}$	$19.95^{+0.35}_{-0.40}$	$17.25^{+0.23}_{-0.22}$	15.30	12.65	11.10

Note: The confidence intervals for m_x for $i = 45°$, $55°$, and $65°$ are not indicated, since the model of the binary system is rejected at the $\alpha = 5\%$ significance level in these cases.

relativistic object, leading to the maximum distortion of the total absorption profiles of the optical star. As a consequence, the shape of the light curve is more sensitive to variations in the orbital inclination when $q < 1$ than when $q > 1$.

The size of the projection of the optical star onto the plane of the sky varies during its orbital motion. The degree of asymmetry of the absorption profiles depends not only on the asymmetry of the velocity field of the area elements and the distributions of $T_{\rm loc}$ and $\log g_{\rm loc}$ relative to the geometrical center of the projection, but also on gradients of these parameters over the surface of the optical component. The most extreme region from the point of view of its physical parameters and their gradients is the nose of the optical star (which is close to filling its Roche lobe). At orbital phases 0.35–0.45 and 0.55–0.65, a large fraction of this nose is visible to the observer. Therefore, the distributions of $T_{\rm loc}$ and $\log g_{\rm loc}$ and their gradients are maximally asymmetric relative to the geometrical center of the projected star. This makes the integrated absorption profiles of the optical star maximally asymmetric, and hence maximally sensitive to variations in the orbital inclination, at orbital phases 0.35–0.45 and 0.55–0.65 (Figs. 4 and 5).

Thus, the strongest orbital variability in the absorption profiles in the spectrum of the optical star is observed at orbital inclinations $i \simeq 90°$ and for component-mass ratios $q < 1$, since, in this case, the nose of the optical star is eclipsed by the star's body

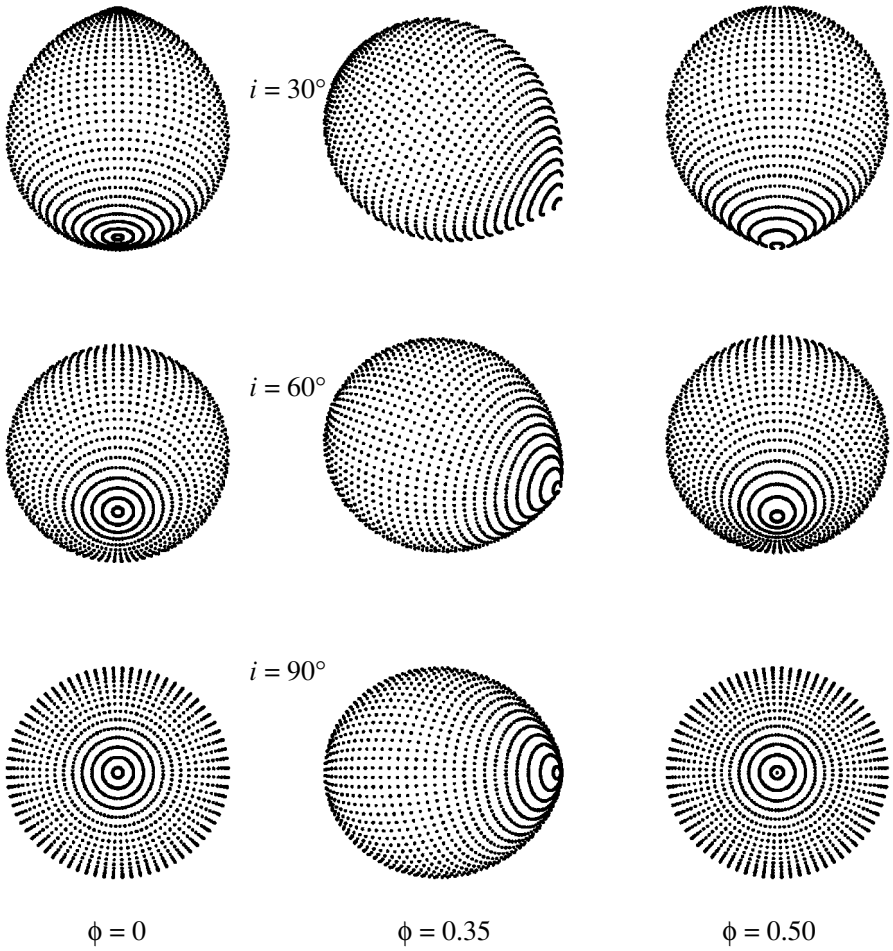

Fig. 12. Optical star in an X-ray binary in a Roche model with $q = 1$, a degree of filling of the Roche lobe $\mu = 1$, and orbital inclinations $i = 30°, 60°, 90°$ at orbital phases 0.0, 0.35, 0.5. When $i = 90°$, the projection of the star onto the plane of the sky varies strongly (from being circular to pear-shaped). The "nose" of the star is sometimes eclipsed and sometimes not, leading to strong distortion of the line profiles and radial-velocity curve. When $i = 30°$, the shape of the projection of the star onto the plane of the sky is nearly constant and varies only with the position angle of the figure, which does not influence the shape of the radial-velocity curve.

during the orbital motion. The shape of the star's projection onto the plane of the sky varies strongly when $i = 90°$, from being circular to pear-shaped, which gives rise to the maximum asymmetry in the radial-velocity curve. When $i \ll 90°$, the nose of the optical star is not eclipsed during the orbital motion. The projections onto the plane of the sky of the shape of the star and of regions on its surface remain nearly constant (only the orientation of the stellar shape varies), making the line profiles less sensitive to the orbital shifts of the optical star. This is illustrated in Fig. 12.

6. CONCLUSION

The main result of this study is our qualitative estimate of the variations of the radial-velocity curve of the optical star in a close binary system as a function of the orbital inclination i and component-mass ratio q for stars with various masses and temperatures (Tables 3–6).

Another important result is our new fitting of an accurate radial-velocity curve for the optical star in the Cyg X-1 binary system using a new and more correct algorithm to synthesize the theoretical radial-velocity curve (based on calculating model atmospheres for each area element [5, 6]). The computations show that the accuracy of the radial-velocity curve in this case ($\sim$3% of the radial-velocity half-amplitude K_v) is sufficient to place constraints on the orbital inclination, using only one radial-velocity curve: $i < 43°$.

The results of test computations show that the variation in the radial-velocity curve due to variations in the orbital inclination i grows with the mass of the optical star; this effect is greatest in binary systems with component-mass ratios $q < 1$ (Tables 3–6). We can adopt the variations in the radial-velocity curves presented in Tables 3–6 to estimate the accuracy of observed radial-velocity curves required to reliably derive the orbital inclinations of observed binary systems. Inclination estimates are most feasible for X-ray binaries with massive optical components ($m_v > 10\,M_\odot$ for systems whose relativistic component is a black hole with mass $m_x \simeq 10\,M_\odot$). The required accuracy of the observed radial-velocity curve should be better than 4–8%. Note that all our computations were carried out for a mass of the relativistic object of $m_x = 10\,M_\odot$. Since the variations in the radial-velocity curve are determined primarily by the ratio $q = m_x/m_v$, our main conclusions are also applicable for X-ray binary systems with neutron stars with masses $m_x \simeq 1.4\,M_\odot$. In this case, the magnitude of the radial-velocity variations could be sufficient to estimate the orbital inclination if the mass of the optical star is $m_v \gtrsim 1\,M_\odot$.

Recall that we have synthesized the radial-velocity curves using the Hγ and CaI 6439 Å absorption lines assuming LTE. It is known that the approximation of LTE is not suitable for hot stars with $T_{\rm eff} > 20\,000$–$30\,000$ K [11]. The inferred radial-velocity variations with variations in the orbital inclination could change when the radial-velocity curve is synthesized allowing for non-LTE corrections. However, note that our test computations using a constant profile shape for the Hγ absorption line over the stellar surface (Figs. 1, 2) show that the variations in the shape of the integrated line profile are mainly sensitive to the geometry of the shape of the star (the position of the star's "nose" relative to the barycenter of the binary system). This provides hope that, although allowance for non-LTE effects could appreciably affect the equivalent widths of the Hγ line in the case of high $T_{\rm eff}$, the variations of the shape of the integrated Hγ profile with orbital phase will differ only slightly with the use of LTE and non-LTE model atmospheres for the optical star. We plan to investigate this question further by carrying out corresponding computations in a non-LTE approximation.

ACKNOWLEDGMENTS

The authors thank L.I. Mashonkin for valuable comments. This work was supported by the Russian Foundation for Basic Research (project no. 02-02-17524) and the program "Leading Scientific Schools of Russia" (grant no. NSh-388.2003.2).

REFERENCES

1. S. Sofia and R. E. Wilson, Astrophys. J. **203**, 182 (1976).
2. É. A. Antokhina and A. M. Cherepashchuk, Pis'ma Astron. Zh. **23**, 889 (1997) [Astron. Lett. **23**, 773 (1997)].
3. T. Shahbaz, Mon. Not. R. Astron. Soc. **298**, 153 (1998).
4. M. K. Abubekerov, É. A. Antokhina, and A. M. Cherepashchuk, Astron. Zh. **81**, 606 (2004) [Astron. Rep. **48**, 550 (2004)].
5. É. A. Antokhina, A. M. Cherepashchuk, and V. V. Shimanskii, Izv. Ross. Akad. Nauk, Ser. Fiz. **67** (3), 293 (2003).
6. É. A. Antokhina, A. M. Cherepashchuk, and V. V. Shimanskii, Astron. Zh. **82**, 131 (2005) [Astron. Rep. **49**, 109 (2005)].
7. M. K. Abubekerov, É. A. Antokhina, and A. M. Cherepashchuk, Astron. Zh. **81**, 108 (2004) [Astron. Rep. **48**, 89 (2004)].
8. É. A. Antokhina and A. M. Cherepashchuk, Astron. Zh. **71**, 420 (1994) [Astron. Rep. **38**, 367 (1994)].
9. R. L. Kurucz, CD-ROMs (Smithsonian Astrophysical Observatory, Cambridge, 1992).
10. A. M. Cherepashchuk, N. A. Katysheva, T. S. Khruzina, and C. Yu. Shugarov, *Highly Evolved Close Binary Stars: Catalog* (Gordon and Breach, New York, 1996), Vol. 1, Part 1, p. 82.
11. A. Herrero, *IAU Symp. No. 212: A Massive Star Odyssei: from Main Sequence to Supernova,* Ed. by Karel A. van der Hucht, Artemio Herrero, and Cesar Esteban (Astron. Soc. Pac., San Francisco, 2003), p. 3.
12. V. Straizhis, *Metal-Deficient Stars* (Mokslas, Vilnius, 1982) [in Russian].
13. É. A. Antokhina, Astron. Zh. **73**, 532 (1996) [Astron. Rep. **40**, 483 (1996)].
14. N. I. Balog, A. V. Goncharskii, and A. M. Cherepashchuk, Pis'ma Astron. Zh. **7**, 605 (1981) [Sov. Astron. Lett. **7**, 336 (1981)].

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 814–825. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 914–927.

Three-Dimensional Hydrodynamical Modeling of the Formation of the Accretion Disk in the SS 433 Binary System

V. V. Nazarenko, L. V. Glazunova, and S. V. Nazarenko

Astronomical Observatory, National University, Shevchenko Park, Odessa, 270014 Ukraine

Received June 6, 2004; in final form, September 20, 2004

Abstract—Three-dimensional hydrodynamical modeling of the formation of the accretion disk in the SS 433 binary system is carried out with various types of cooling and numerical grids. These computations show that a thick accretion disk with a height of 0.25–0.30 (in units of the component separation) is formed around the compact object, from a flow with a large radius (0.2–0.3 in the same units) that forms in the vicinity of the inner Lagrangian point. This disk has the form of a flattened torus. The number of orbits of a particle of gas in the disk is 100–150, testifying to a minimal influence of numerical viscosity in these computations. The computations also show that the stream flowing from L_1 is nearly conservative, and spirals in the disk are not formed due to the influence of the donor gravitation. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The binary system SS 433 (V1343 Aql) is named according to its number in the catalog of objects with Hα emission lines compiled by Stephenson and Sanduleak [1]. Margon *et al.* [2] made the sensational report that three groups of Hα and neutral He emission lines in the spectrum of the star displayed huge shifts of about 1000 Å relative to their normal positions. This was explained by the presence of narrow gaseous jets moving perpendicular to the plane of an accretion disk at velocities of 80 000 km/s [3–6]. The jets precess with a period of about 164 days, and the angle between the jet axis and the rotational axis of the star is roughly 20°. SS 433 has been actively studied over the last 20 years; it is commonly thought to be a close binary system containing a compact object—a neutron star or black hole [4, 7, 8].

The optical eclipses detected for SS 433 in 1981 by Cherepashchuk [9] made it possible to determine the basic parameters of the system. SS 433 is a close binary with an orbital period of 13.086 days [10] consisting of an O–B star and a compact object surrounded by an optically thick accretion disk. A photometric study of the light curve of the system revealed three periods of the object's brightness variations: the 13-day orbital period, 164-day precessional period, and approximately 6-day nutational period [11, 12]. Observations also indicate the presence of nonperiodic variations of the optical continuum brightness on timescales of hours and minutes.

Currently, two basic models of SS 433 are considered, corresponding to a high-mass and low-mass system. In the first case, the mass function of the compact object is either 10.1 $M_\odot$ [13] or 7.7 $M_\odot$ [14]. In the case of a low-mass system, the mass function is equal to 2.0 $M_\odot$ [15]. The different estimates for the mass ratio correspond to the spectral interval for the observations: $q = M_x/M_{\rm opt} = 0.15$–$0.25$ (X-ray) and $q = 1.2$ (optical [10]). It is commonly thought that the optical star overflows its Roche lobe and loses mass at a rate of 7.0×10^{-6}–$4.0 \times 10^{-4}\ M_\odot$/yr [16]. However, contrary to this proposed very high rate of mass transfer through the point L_1, observations do not indicate appreciable variations of the orbital period: the variations of $\Delta P/P$ are smaller than 10^{-6}.

Thus, in spite of very active studies of this system, little is known about the processes occurring in it. There is not even general agreement regarding the system's parameters. Hydrodynamical calculations of the gas motions in the vicinity of the system components may help clarify the situation.

We present here the results of three-dimensional hydrodynamical modeling of the formation of the accretion disk in SS 433.

Since the path of a gas particle between collisions is much shorter than the distance between the centers of the binary components, the hydrodynamical approach to the problem of mass transfer in close binary can be basic. Here, for simplicity of the computations, we will consider the solid medium to be an ideal gas consisting purely of hydrogen, rather than a plasma (as is the case in real binaries).

Two-dimensional hydrodynamical calculations of the gas motions in a close binary and of the formation of the associated accretion disk were carried out by Sawada *et al.* [17] and Matsuda *et al.* [18], who also first obtained a spiral shock wave in the accretion disk. Analytical two-dimensional calculations

1063-7729/05/4910-0814$26.00

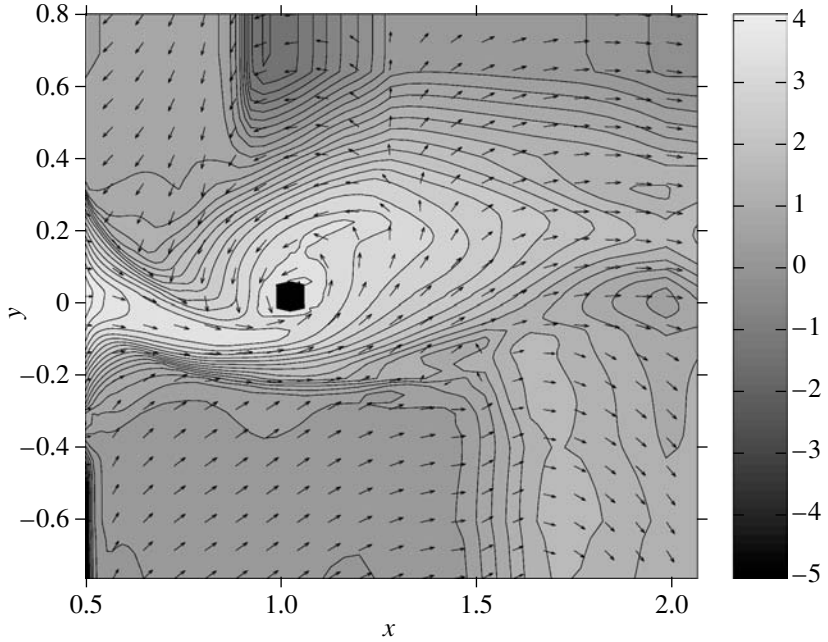

Fig. 1. Density contours and the velocity field for the cross section made by the orbital plane (cooling via regularization of the temperature field).

of spiral shocks were carried out in [19, 20]. Three-dimensional hydrodynamical calculations of the flow from L_1 in the Roche lobe of the accreting star and the formation of the accretion disk were presented in [21–32], using short-period close binaries (cataclysmic variables) as an example. Since SS 433 is substantially different from cataclysmic variables due to its long orbital period, these three-dimensional hydrodynamical calculations cannot be applied directly to this system. Thus, to describe the envelope components in SS 433, computations must be carried out directly using the actual parameters of the system.

There have been only three studies in which the gas motions in SS 433 have been calculated directly. Chakrabarti and Matsuda [33] and Nazarenko *et al.* [34] carried out two-dimensional modeling of the system, while the structure of the flow was studied by Bisikalo *et al.* [35]. Spiral shocks in the accretion disk were calculated in [33], where it was shown that these shocks could plausibly result in variations of the system's brightness on timescales shorter than a day. Unfortunately, the computational domain used in [33] did not cover the entire space between the components, so that the resulting pattern of the gas motions was incomplete. This gap was filled by the study [34], in which the computational domain included the entire space from L_2 to L_3 and from L_4 to L_5 and both components of the system. The following results were obtained in this study. First, the low-mass model for SS 433 is unable to yield a mass-transfer rate through L_1 exceeding $10^{-6}\,M_\odot$/yr, making this model inappropriate as a description of the mass transfer in the system. Second, in order to bring the observed mass-transfer rate through L_1 into agreement with the fact that the relative variations of the period are below 10^{-6}, the ratio of the masses of the compact object and the optical star must exceed 0.9. Third, the flow moving in the Roche lobe of the compact object from L_1 forms an accretion disk and circumstellar envelope, which surrounds the disk and the system components. Approximately 50% of the matter in the flow ends up in the accretion disk, while the remaining 50% forms the envelope, from whose peripheral zone gas leaves the binary system.

One drawback of [34] is that the computations did not include radiative cooling of the moving gas explicitly (instead, the ratio of specific heat capacities γ was simply decreased to a value close to unity, as is usual in such modeling). As a result, the mass exchange in the system becomes nonconservative (without taking radiative cooling into account explicitly, the heating of the moving gas due to numerical viscosity becomes large, resulting in an increase in the temperature and, accordingly, in the role of the gas pressure).

Here, we present the results of three-dimensional hydrodynamical computations of the gas motion in SS 433 taking into account radiative cooling in various ways.

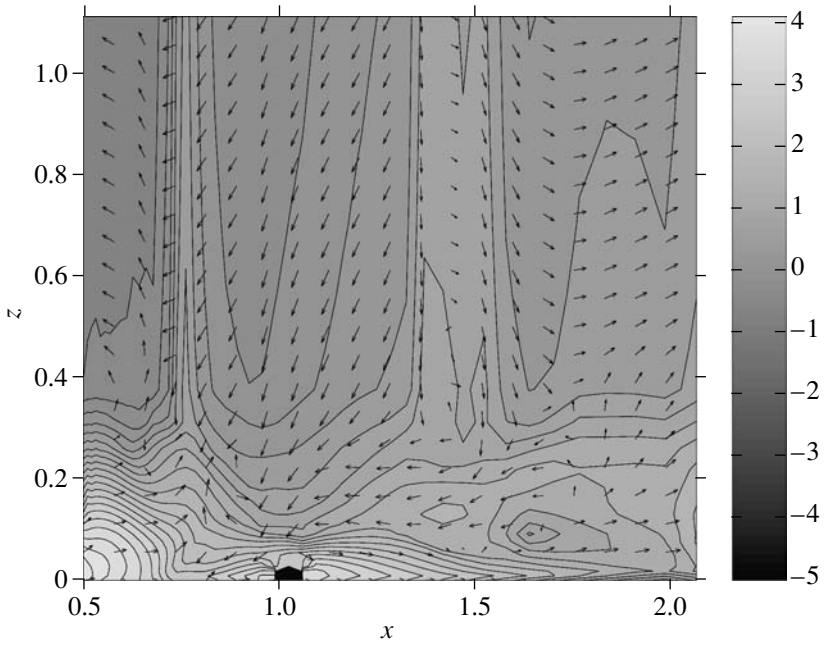

Fig. 2. Same as Fig. 1 for the cross section made by the $z-x$ plane, which contains the line connecting the centers.

2. MODEL FOR THE MASS TRANSFER: INITIAL AND BOUNDARY CONDITIONS

We use here the same approach as in our studies [34, 36, 37]. Briefly, it consists of the following: (1) the formation of the flow from the donor atmosphere through L_1 is calculated to determine the initial parameters of the flow (its structure in the vicinity of L_1), and (2) the initial configuration of the donor atmosphere in the vicinity of L_1 is reconstructed using Kurucz model atmospheres [38].

This way of determining the initial parameters of the flow is important for close binaries in which the region of overflow of the Roche lobe surrounding the donor exceeds the size of its atmosphere. SS 433 is such a system: to provide the required high mass-transfer rate, the point L_1 must be located in very deep inner layers of the donor [34].

To solve the nonstationary hydrodynamical equations in Euler form, we applied a version of the "method of large particles" developed by Belotserkovsky and Davydov [39], in which artificial viscosity is specified in the first stage of calculations and which has second order accuracy in space and time.

As one of gas cooling techniques, as well as to make the calculations stable, we applied a special technique, which is used in this version of the method of "large particles," at the second stage of each time step, in those cells where the temperature began to exceed some limiting value $T_{\rm lim}$.

The essence of this technique, which we call temperature-field regularization, is that, if the temperature in a cell exceeds $T_{\rm lim}$, it is set equal to the value in the first stage of the calculations for the given time step in this cell. This corresponds to a sort of cooling, which should restrict the effects of heating of the moving medium by numerical viscosity. In the method of large particles, the parameters of the flow in a given cell at each time step are determined in two stages: in the first, the effects of gas pressure and external fields (in our case, gravitation) are calculated, while, in the second, the transport of mass, momentum, and internal energy through the cell boundaries is calculated taking into account the velocities obtained in the first stage, and the final values for the physical parameters at the given time step in a given cell are determined.

Our computations were carried out in a rotating coordinate frame with its origin at the center of the donor and with the x, y, and z axes forming a right-handed set. The computational domain (in units of the distance between the components) occupies the volume of space from L_1 ($x = 0.456$) to $x = 1.95$, from $y = -1.45$ to $y = 1.45$, and from the orbital plane ($z = 0$) to $z = 0.56$. Thus, the libration points L_1, L_3, L_4, and L_5 are all within the computational domain, making it possible to trace the entire path of the gas flowing from L_1 within the accretor's Roche lobe. Since the computational domain contains the entire accretion disk, a complete three-dimensional model of the disk can be constructed.

The computations were carried out for two three-dimensional numerical grids: a high-resolution grid with $140 \times 140 \times 70$ cells, and a low-resolution grid with $90 \times 90 \times 50$ cells. The high-resolution grid was

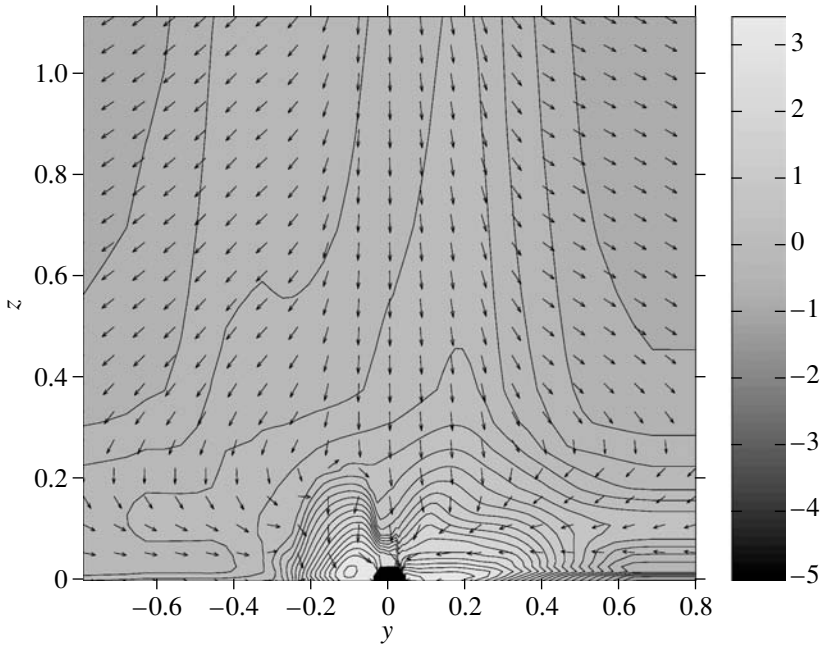

Fig. 3. Same as Fig. 1 for the cross section made by the $z-y$ plane, which contains the center of the accretor.

used to represent a model of the disk with low viscosity, and the low-resolution grid a model with high viscosity.

We used a standard model for the binary system: the rotation of the components about their axes and their orbital motion were assumed to be synchronous, the gravitational fields of the close-binary components were treated in a Roche approximation, all effects of the rotation (centrifugal and Coriolis forces) were included, and the orbits of the components were assumed to be circular.

In the computations, we adopted for SS 433 a donor mass of 10 $M_\odot$, an accretor mass of 12 $M_\odot$, an orbital period 13.086 days, and an effective temperature for the donor of 22 500 K.

3. RESULTS

3.1. Cooling via Temperature-Field Regularization

The degree of Roche lobe overflow in the computations was 0.06, which corresponds to a mass-transfer rate through L_1 of $3.0 \times 10^{-5}\ M_\odot$/yr.

The computations were made over 12 orbital periods in order to reach a steady state in the computational domain over a long time. We took a steady state to be reached when the mass-transfer rate through L_1 was equal to the accretion rate onto the disk and the mass-loss rate through the outer boundaries of the computational domain. The steady state in the accretion disk was reached after approximately two or three orbital periods. In the vicinity of L_1, steady state was reached after approximately one-sixth of the orbital period.

Figures 1–3 present the results of the computations in the form of contours of constant density and the velocity field. Figure 1 presents this information in the cross section of the computational domain made by the orbital plane (all the density contours are given in units of 10^{11} cm^{-3}, and the velocity field is given in arbitrary units). We can see the flow emerging from the vicinity of L_1, as well as the accretion disk that has formed from this flow. The width of the flow is 0.2–0.3, or 50% of the donor radius. It is difficult to determine the initial radius of the flow more accurately, since the density in the direction perpendicular to the flow axis does not decrease to small values, even at the boundary of the computational domain, due to the presence of the common envelope surrounding the disk and donor. As we can see from Fig. 1, the width of the flow remains approximately the same along its entire length. The radius of the disk is 0.34–0.4, which is 60–70% of the accretor's Roche lobe. The density of the flow in the vicinity of the accretor is 10^{14} cm^{-3} and increases to 10^{13}–10^{14} cm^{-3} in the disk. The temperature in the flow is 3.0×10^4 K, and in the disk close to its outer boundary, it is 3.5×10^4 K. The temperature in the inner parts of the disk is unrealistically high, of the order of 10^6 K (apparently because the cells are too large in the region close to the accretor, where the gravitational field varies dramatically, resulting in substantial heating due to numerical viscosity). The formation of a disk without a hot spot, i.e., in the absence of a direct interaction between the disk and flow,

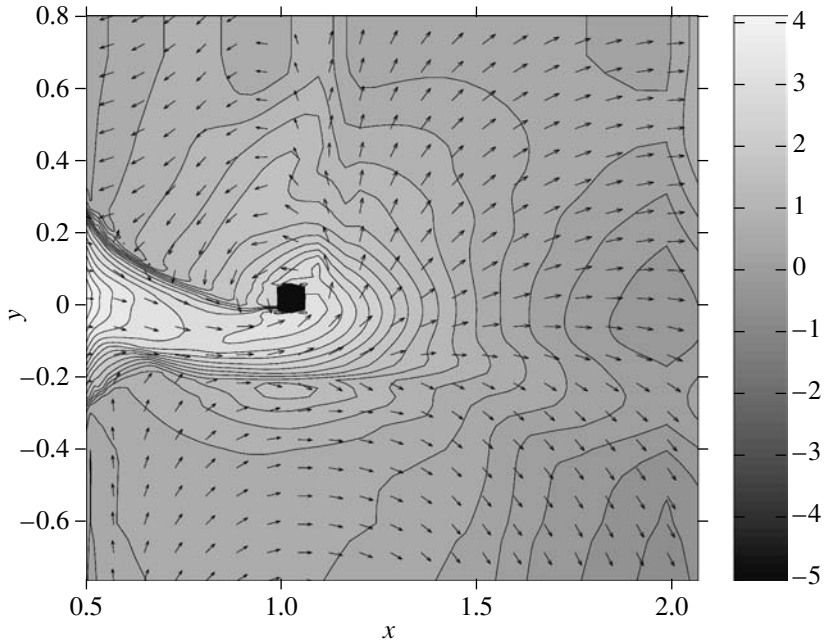

Fig. 4. Density contours and the velocity field for the cross section made by the orbital plane (with radiative cooling taken into account implicitly, $\gamma = 1.0001$).

is due to the fact that the specific angular momentum of the gas particles relative to the accretor decreases along their trajectory from the flow to the disk; thus, the specific angular momentum of the gas in the flow exceeds that in the disk. Instead of a hot spot, the boundary of the flow in the direction of orbital phases 0.15–0.3 displays a so-called "hot line" (the term used by Bisikalo *et al.* [30]). In our version of radiative cooling, the disk does not contain any spiral shocks or structural peculiarities; i.e., it displays a very smooth structure, apparently because insufficient cooling of the gas leads to smoothing due to the resulting high gas pressure.

Estimates of the amount of matter lost from the system through its outer boundaries indicate that the flow from the inner Lagrangian point is essentially conservative, with only about 1% of the matter being lost.

The height of the flow in the z direction is roughly 0.2, which is essentially equal to the hydrostatic height of the flow at the given point.

Figures 2 and 3 present density contours and the velocity field in cross sections of the computational domain made by the $z-x$ plane, which contains the line connecting the centers, and $z-y$ plane, which contains the center of the accretor. These figures show that, in the z direction over the entire surface of the disk, the disk height varies within 0.15–0.35 and the disk displays a flattened toroidal shape. The only exception is in the inner part of the disk, close to the accretor, where the disk thickness decreases to approximately hydrostatic heights, which are equal to 0.02–0.05. Thus, the thickness of the disk in the system is close to the initial thickness of the flow. The substantial overflow of the Roche lobe by the donor results in the formation of a flow with a large radius and ultimately of a thick, opaque disk. With this type of radiative cooling, the temperature is approximately the same over the entire computational domain beyond the orbital plane and is equal to 35 000–45 000 K.

3.2. Computation of the Flow Motion and Formation of the Accretion Disk with Implicit Radiative Cooling

We specified radiative cooling implicitly in the computations by decreasing the ratio of the specific heat capacities γ, using the low-resolution grid. Figures 4–6 illustrate the mass transfer from the inner Lagrangian point and the formation of the accretion disk in the system for $\gamma = 1.0001$. We can see from Fig. 4 that, with this type of cooling, the flow formed in the vicinity of L_1 also has a radius of 0.2–0.3 (this is natural, since the degree of overflow of the Roche lobe by the donor and the system parameters remained unchanged). The flow in the accretor's Roche lobe displays a sharp boundary. This is due to the physical conditions in the donor atmosphere, from which the flow has been formed, namely, the exponential decrease in the density at the outer boundary of the atmosphere. The width of the flow does not vary significantly and only displays a slight decrease in the vicinity of the accretor. Figure 4

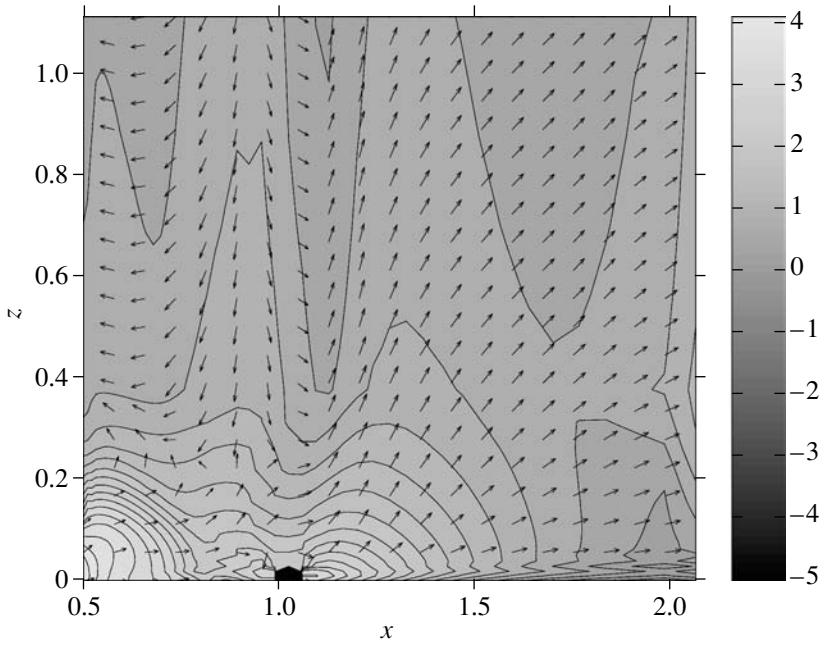

Fig. 5. Same as Fig. 4 for the cross section made by the $z-x$ plane, which contains the line connecting the centers.

shows that the shape of the disk formed by the flow is almost circular in the orbital plane. The density of the matter in the flow is roughly 10^{14} cm^{-3} in the vicinity of the accretor and $10^{13}-10^{14}$ cm^{-3} in the disk. The disk displays an essentially smooth structure, without any distinct features, apparently due to insufficient cooling of the gas (the high gas pressure associated with the numerical viscosity smooths nearly all details of the structure). Only at the outer boundary of the disk facing orbital phases 0.25–0.3 can we see an indistinct spiral shock (see the density contours in Fig. 4) formed by the interaction between the gas from the disk and the gas from the outer envelope with density 10^{11} cm^{-3}. Such a low density in the outer envelope is due to the conservativeness of the mass exchange in the computations. The flow displays a comparatively low temperature in the vicinity of L_1, 75 000 K approximately. This value corresponds to a high degree of overflow of the Roche lobe by the donor. The temperature increases to 100 000 K with approach toward the accretor. The temperature in the disk varies from 50 000 K (at the outer boundary of the disk facing orbital phase 0.25) to approximately 10^6 K (in the inner parts of the disk, close to the accretor). We can see the hot line formed when gas from the disk interacts with the outer boundary of the flow facing orbital phases 0.15–0.3, where the temperature reaches approximately 10^5 K.

Figures 5 and 6 present plots in cross sections of the computational domain made by the $z-x$ and $z-y$ planes, which contain the line connecting the centers and the line perpendicular to this passing through the accretor. In these planes, the disk displays an almost cylindrical shape in the vertical direction, with a height equal to 0.24–0.34, substantially exceeding the hydrostatic height, which is roughly 0.12–0.17. Thus, the formation of the thick, opaque disk can be explained by the substantial initial height of the flow emerging from L_1. Only in the vicinity of the accretor is the height of the disk close to the hydrostatic value, roughly 0.05–0.07.

The flow also displays a sharp boundary in the z direction, where its height is about 0.22, almost equal to the hydrostatic height at the given point (roughly 0.20).

3.3. Computations of the Flow Motion and the Formation of the Accretion Disk with Combined Radiative Cooling and the High-Resolution Numerical Grid

The "combined" radiative cooling simultaneously includes implicit radiative cooling via a decrease in the ratio of specific heat capacities, and cooling via regularization of the temperature field. The mass-transfer rate through L_1 is $1.2 \times 10^{-5}\, M_\odot$/yr, which corresponds to a degree of overflow of the Roche lobe by the donor of 0.05. In this case, at L_1, the density is 1.3×10^{14} cm^{-3}, the temperature is 85 000 K, and the ratio of specific heat capacities $\gamma = 1.0001$.

Figure 7 presents the structure of the disk (density contours and stream lines) in the orbital plane at the end of the computations. The width of the flow in the vicinity of L_1 is roughly 0.15–0.20. The flow

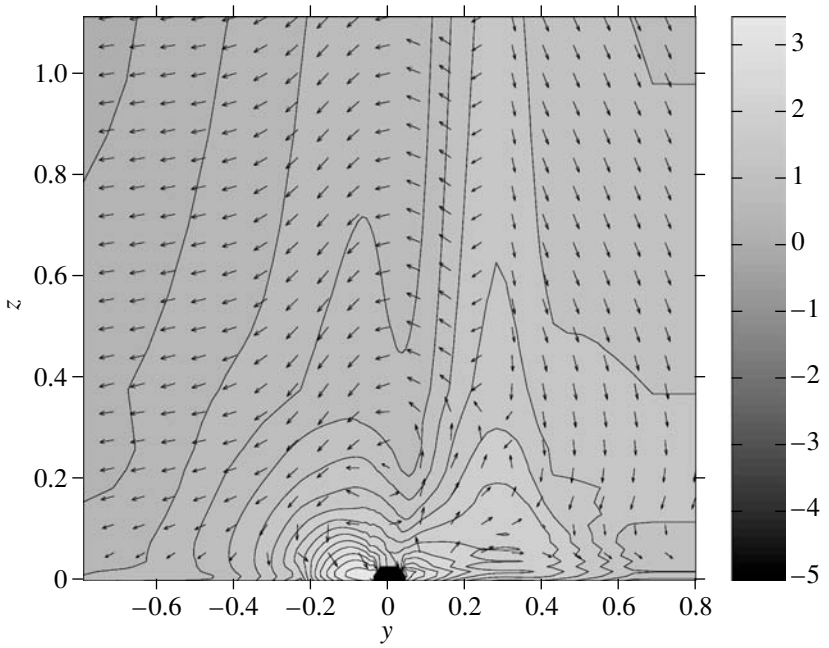

Fig. 6. Same as Fig. 4 for the cross section made by the $z-y$ plane, which contains the center of the accretor.

displays essentially the same radial size as it moves in the Roche lobe of the accretor, about 0.20. The size of the flow in the z direction at L_1 is close to the size in the orbital plane, 0.12–0.17. In the course of its subsequent motion in the Roche lobe, the size of the flow in the z direction decreases to 0.09–0.14, providing evidence for some constriction of the flow in this direction.

Near the accretor, the flow forms the accretion disk, in which a particle makes approximately 90–150 revolutions around the accretor (this value is somewhat different at different times, probably reflecting variations in the viscosity in the disk after the steady state is established). This number of revolutions indicates that the disk is elliptical, with a fairly low viscosity, and is realistic (i.e., the number of revolutions corresponds to the number derived analytically based on two parameters of the disk—the accretion rate and the height in the z direction). Two spiral shocks (I and II) are formed in the disk, which unwind in the direction opposite to the binary rotation, from regions close to the accretor to the outer edge of the disk. We can also see a short spiral shock III that is wound in the direction of the binary rotation (the leading shock; see, for example, [40]). In addition, the edge of the disk facing the flow displays a perturbation formed during the interaction between the flow and disk—the so-called hot line (denoted IV in Fig. 7).

Thus, the disk displays a complex inner structure and forms without a hot spot: the matter flows into the disk along a tangent, without any direct impact. Initially, the matter from the flow forms a "halo" around the disk [41], with a large fraction of the gas from this halo subsequently falling onto the disk (around 99%), while the remaining gas leaves the system. This distribution of matter in the computational domain indicates that the flow is essentially conservative. However, final conclusions regarding the conservativeness of the mass transfer in SS 433 can only be obtained using a numerical grid that covers the entire area surrounding the system's components.

The fact that, in our calculations, as in those of Bisikalo *et al.* [31, 41] for a cataclysmic system, no hot spot is formed during the interaction of the flow and disk suggests that this is typical of binaries in general.

To investigate the origin of the spiral shock waves, we eliminated the gravitational field of the donor from the region of the disk at some stage of the computations (after 9 orbital periods). This made it possible to check whether the outflowing star affects the formation of the spiral wave. Figure 8 displays the structure of the disk in the orbital plane with the gravitational field of the donor switched off. We can see from Fig. 8a that, after the donor's gravitation is removed in the disk (this was done within a distance that did not exceed 0.45 from the disk center; at that time, the radius of the disk was about 0.35), the disk is constricted in the orbital plane, and its radius decreases from 0.30–0.35 to 0.17–0.20 (by 30%). The spirals in the disk maintain their positions, but their shapes become somewhat indistinct, and they are only slightly visible in the density contours. In

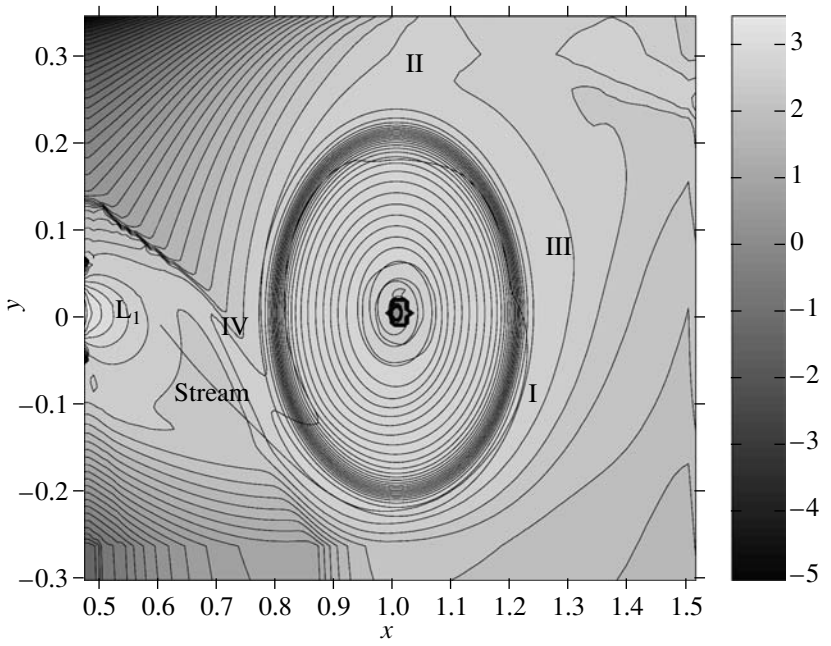

Fig. 7. Density contours and stream lines in the disk in the cross section made by the orbital plane (combined radiative cooling and high-resolution grid).

contrast, in Fig. 8b, which presents the corresponding temperature contours, the spiral shocks I, II, III are much more distinctly visible than they were before the donor's gravitation was switched off. Thus, the origin of the spiral shocks in the disk remains unclear.

Strange as it may seem, the hot line essentially disappears when the donor's gravitation is switched off in the disk; this feature is only slightly visible in the temperature contours. This seems to be independent of the donor's gravitation and its role in the formation of the hot line, and is simply due to the fact that the radius and shape of the disk have changed substantially in this case.

One result that stands out is that, after switching off the donor gravitation, the number of revolutions of a particle about the accretor decreases dramatically (to five revolutions); i.e., the disk's viscosity has become much higher.

To verify that the gravitation of the donor does not affect the formation of the spiral structure, we have also calculated the formation of the accretion disk with the gravitation field of the donor switched off within a sphere with radius 0.42 right from the very start of the computations (note that this sphere contains the entire disk and surrounding region). The spirals form in the disk from the very start of the computations; however, their shapes become very indistinct with time, although they are still visible in the disk structure (see Figs. 9a and 9b, which correspond to different times).

Since our computational domain was restricted to the region of the accretion disk, this proves that the spirals cannot be formed by the influence of the circumstellar envelope on the disk [31]. They are apparently formed by some other internal processes. It may be that the spirals are formed as a result of deformations of the elliptical disk due to its differential rotation, in the same way as spirals in a galactic disk [42].

Figures 10a and 10b present the structure (density contours) of the disk in the $z-x$ and $z-y$ planes. The disk shape is close to cylindrical in these planes; its height is approximately 0.1, substantially smaller than the height obtained for the computations with the coarser grid.

4. CONCLUSIONS

We have modeled the three-dimensional structure of the accretion disk in the SS 433 system assuming $M_d = 10\,M_\odot$, $M_x = 12\,M_\odot$, and a degree of overflow of the donor Roche lobe of 0.06, which corresponds to a mass-transfer rate through L_1 of $3.0 \times 10^{-5}\,M_\odot$/yr. The computations were carried out for three types of radiative cooling: explicit cooling with regularization of the temperature field; standard cooling, in which the ratio of the specific heat capacities is decreased to near unity ($\gamma = 1.0001$); and combined cooling, in which radiative cooling was taken into account implicitly ($\gamma = 1.0001$) and temperature regularization was applied on a high-resolution grid.

Our computations indicate that, in the first two types of radiative cooling (with a coarse grid), the flow from L_1 forms an accretion disk with the approximate

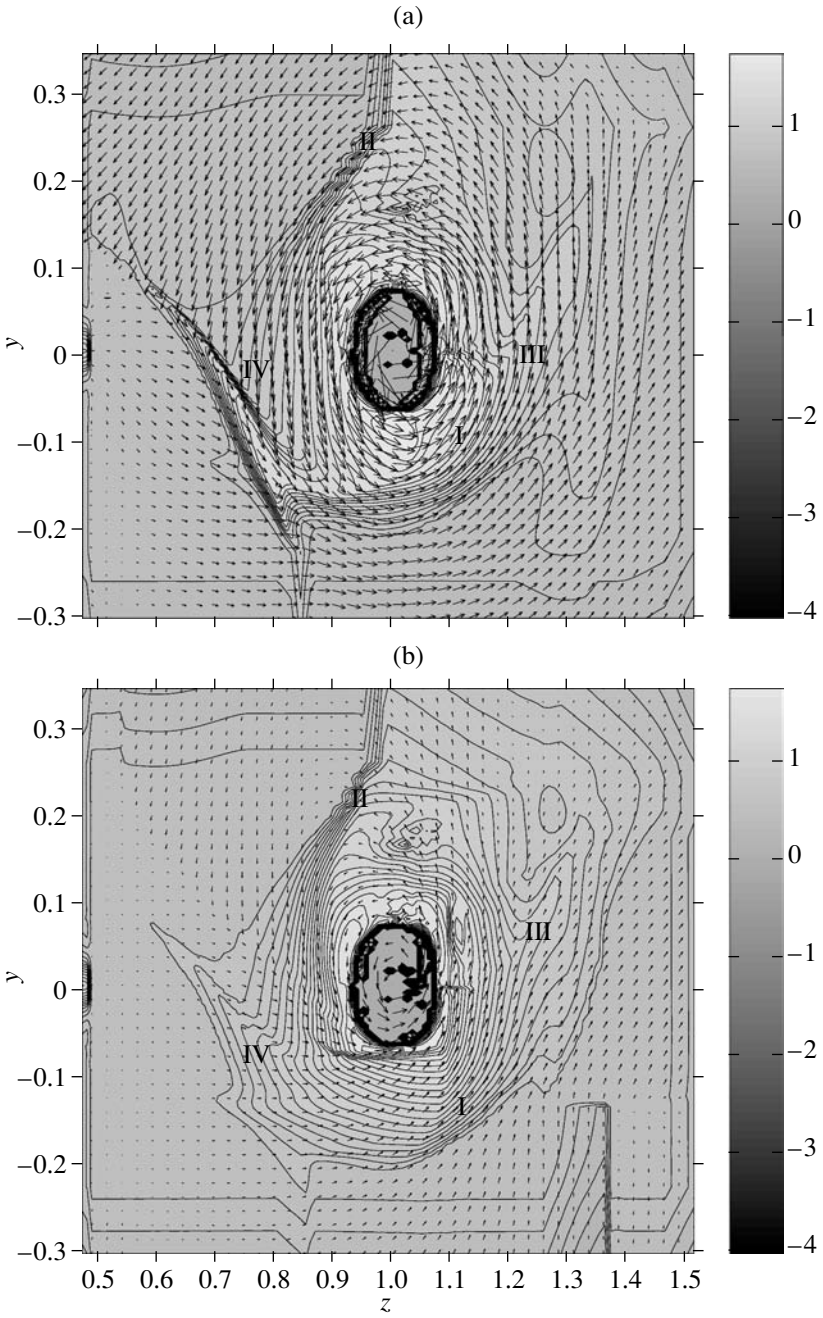

Fig. 8. (a) Density contours and stream lines and (b) temperature contours for the cross section made by the orbital plane (combined radiative cooling without the donor gravitation).

radius 0.34–0.38, which is 70% of the radius of the accretor's Roche lobe. The height of the disk in the z direction is about 0.24–0.34, which is twice the hydrostatic scale height. Thus, the height of the disk is comparable to the initial radius of the flow, and the formation of a thick, opaque disk in SS 433 is probably due to the large initial radius of the flow, which, in turn, results from the extended atmosphere of the star, which appreciably overflows its Roche lobe.

The disk has a shape of a flattened torus; the density of the matter in it is 10^{13}–10^{14} cm^{-3} for a mass-transfer rate through L_1 of $3.0 \times 10^{-5}\,M_\odot$/yr.

Taking into account only the gas flowing from L_1, the mass transfer in the system is essentially conservative. In total, the binary loses only about 1% of the matter flowing from L_1.

The outer boundary of the disk next to the flow displays a hot line, while the other outer boundary of the disk, facing orbital phases 0.25–0.35, displays an indistinct spiral shock. This smearing of the shock is

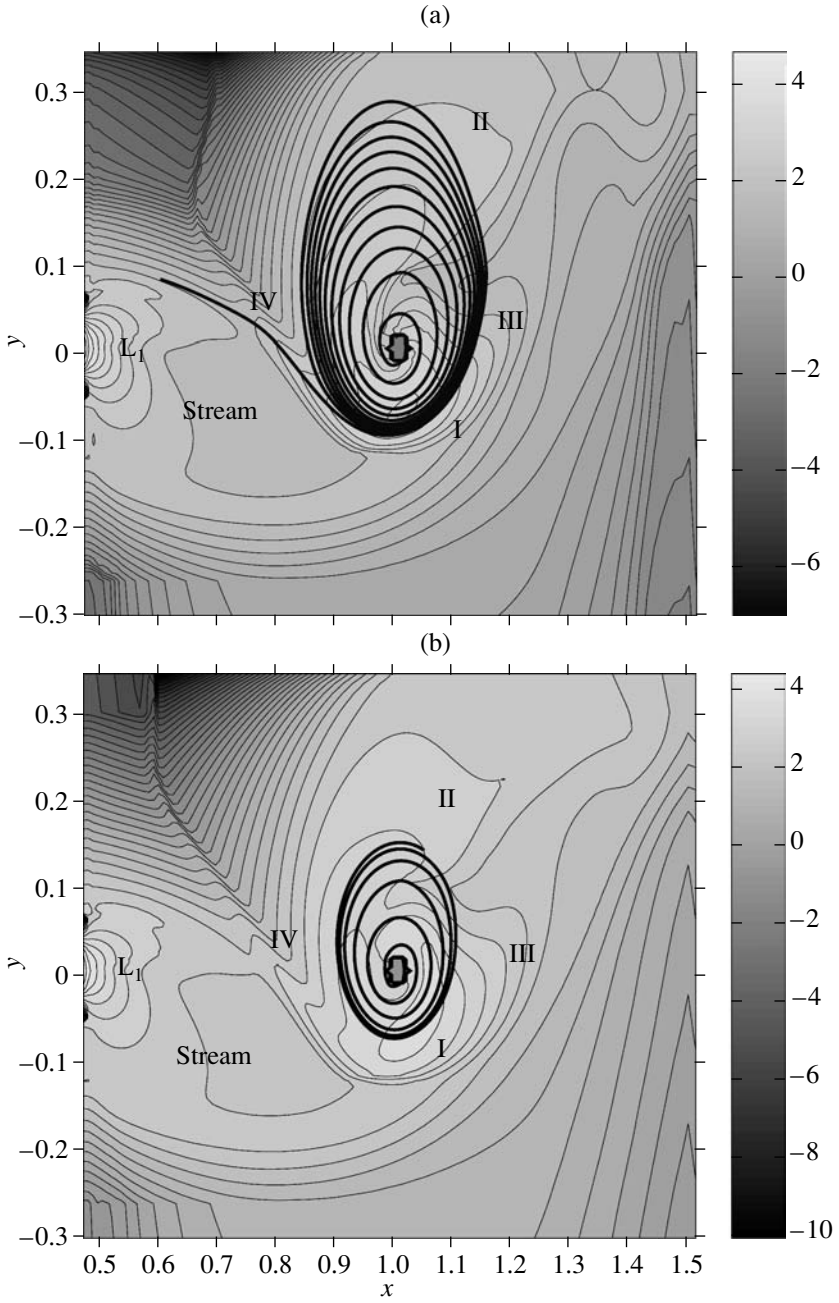

Fig. 9. Density contours and stream lines for the cross section made by the orbital plane (combined radiative cooling). The effect of the donor's gravitation on the formation of the accretion disk was excluded throughout the computations. Two times are presented: (a) 2.24 and (b) 21.24.

apparently due to insufficient cooling of the gas in the computations (too high a temperature in the disk and, especially, in the spirals).

According to our computations with combined radiative cooling using a high-resolution numerical grid, the flow forms an elliptical disk with an average radius of about 0.3. The number of revolutions of a particle in the disk is from 90 to 150 at different times, which is close to the number obtained analytically (around 100) based on calculations of the radial velocity in the disk for the given accretion rate and height of the disk in the z direction. Thus, a fairly realistic disk is formed in these computations, with the number of revolutions occurring in it corresponding well to the disk parameters (its height and accretion rate). Since the number of revolutions of a particle in the disk reflects its viscosity, this indicates that the viscosity is reduced to a minimum in the computations with the high-resolution grid. The number of revolutions does not remain constant even when the steady-state solution has been achieved, providing evidence for variability of the viscosity in

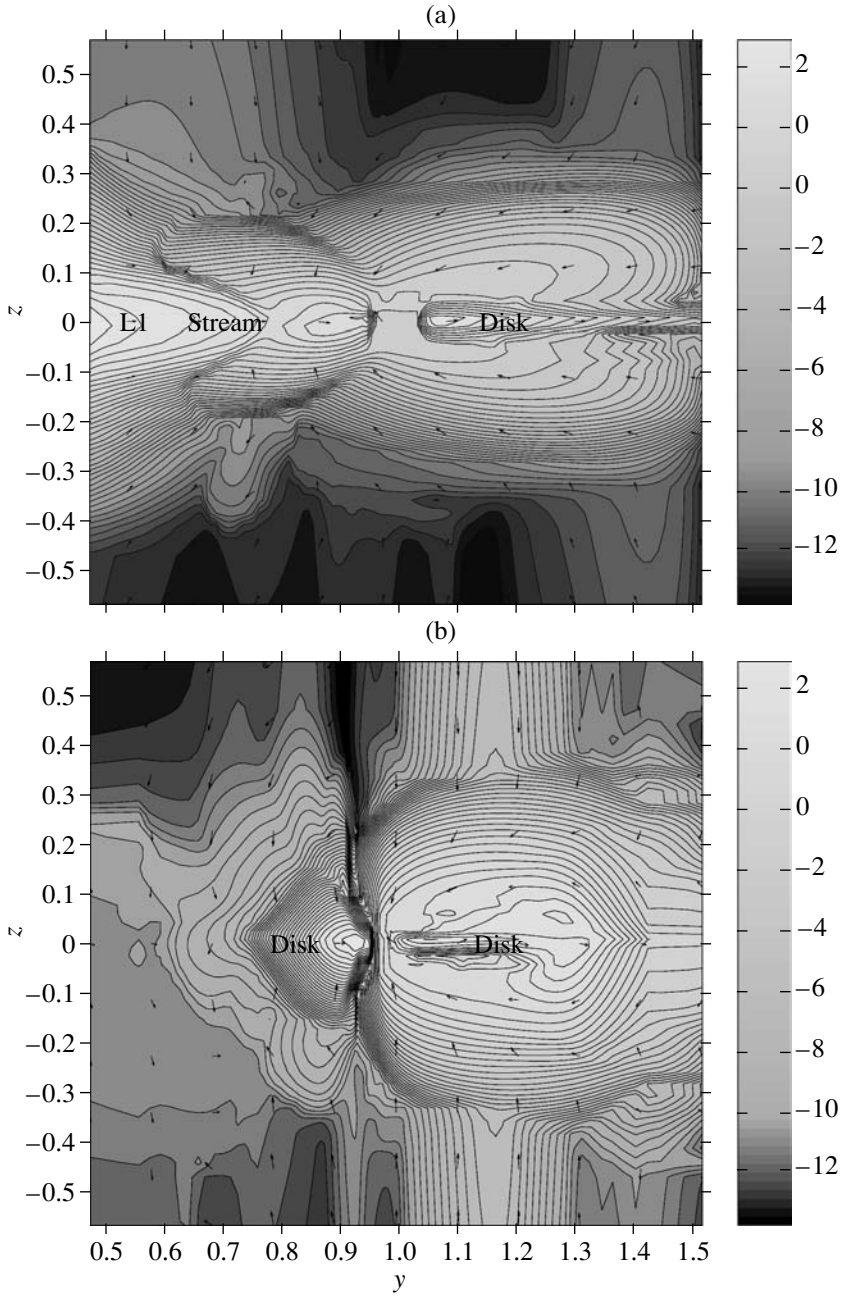

Fig. 10. Density contours for the cross section made by (a) the $z-x$ plane, which contains the line connecting the centers, and (b) the $z-y$ plane, which contains the accretor (combined radiative cooling).

the disk (in deformations of the elliptical disk, the velocity field in the disk varies with time even for the steady-state solution, which apparently affects the disk's viscosity).

Spiral shocks are formed in the disk, which do not disappear even when the gravitational field of the donor is switched off. When this field is switched off, the spiral shocks become somewhat less distinct in the density contours, but remain visible in the temperature contours. Since the outer envelope is excluded from our computations (the computational domain is restricted to the region of the accretion disk), we conclude that the spiral shocks in the disk have some internal rather than external origin. Shocks in the disks of close binary systems may have the same origin as spiral waves in galactic disks. Since our computations indicate that spirals in the disk are independent of the influence of the donor, it may be that a similar spiral structure forms in accretion disks surrounding active galactic nuclei. It is particularly interesting that, when the gravitation of the donor is switched off, the disk decreases slightly in radius and its shape becomes more elliptical. The number of revolutions of a gas particle about the accretor

decreases dramatically; i.e., the disk becomes substantially more viscous. Thus, the donor's gravitation favors the formation of a disk in which gas particles slowly fall towards the disk center (the gravitation of the donor affects the velocity field in the disk, which in turn can influence the viscosity).

Overall, our computations with a high-resolution grid (and, accordingly, low viscosity in the disk) indicate that, to model the formation of more realistic gaseous structures in close binaries, numerical methods with third to fourth order in accuracy and the Navier–Stokes equation should be used together with various physical models for the viscosity, making it possible to eliminate the effect of numerical viscosity in the computations.

REFERENCES

1. C. B. Stephenson and N. Sanduleak, Astrophys. J., Suppl. Ser. **33**, 459 (1977).
2. B. Margon, H. C. Ford, J. I. Katz, *et al.*, Mon. Not. R. Astron. Soc. **189**, 19 (1979).
3. A. S. Fabian and M. J. Rees, Mon. Not. R. Astron. Soc. **187**, 13 (1979).
4. P. G. Martin and M. J. Rees, Mon. Not. R. Astron. Soc. **189**, 19 (1979).
5. R. J. Terlevich and J. E. Pringle, Nature **278**, 719 (1979).
6. S. N. Fabrika, Astrophys. Space Phys. Rev. **12**, 1 (2004).
7. I. S. Shklovskii, Pis'ma Astron. Zh. **5**, 644 (1979) [Sov. Astron. Lett. **5**, 344 (1979)].
8. E. P. J. van den Heuvel, J. P. Ostriker, and J. A. Petterson, Astron. Astrophys. **81**, L7 (1980).
9. A. M. Cherepashchuk, Mon. Not. R. Astron. Soc. **194**, 761 (1981).
10. E. A. Antokhina and A. M. Cherepashchuk, Astron. Zh. **64**, 562 (1987) [Sov. Astron. **31**, 295 (1987)].
11. V. P. Goranskii, V. F. Esipov, and A. M. Cherepashchuk, Astron. Zh. **75**, 240 (1998) [Astron. Rep. **42**, 209 (1998)].
12. V. P. Goranskii, V. F. Esipov, and A. M. Cherepashchuk, Astron. Zh. **75**, 383 (1998) [Astron. Rep. **42**, 336 (1998)].
13. D. Crampton and J. B. Hutchings, Astrophys. J. **251**, 604 (1981).
14. S. N. Fabrika and L. V. Bychkova, Astron. Astrophys. **240**, L5 (1990).
15. S. D'Odorico, T. Oosterloo, T. Zwitter, and M. Calvani, Nature **353**, 329 (1991).
16. A. R. King, R. E. Taam, and M. C. Begelman, Astrophys. J. **530**, 25 (2000).
17. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **219**, 75 (1986).
18. T. Matsuda, M. Inoue, K. Sawada, *et al.*, Mon. Not. R. Astron. Soc. **229**, 295 (1987).
19. H. C. Spruit, Astron. Astrophys. **184**, 173 (1987).
20. S. K. Chakrabarti, Astrophys. J. **362**, 406 (1990).
21. D. Molteni, G. Belvedere, and G. Lanzafame, Mon. Not. R. Astron. Soc. **249**, 748 (1991).
22. M. Hirose, Y. Osaki, and S. Mineshige, Publ. Astron. Soc. Jpn. **43**, 809 (1991).
23. G. Lanzafame, G. Belvedere, and D. Molteni, Mon. Not. R. Astron. Soc. **258**, 152 (1992).
24. G. Lanzafame, G. Belvedere, and D. Molteni, Mon. Not. R. Astron. Soc. **263**, 839 (1993).
25. Z. Meglicki, D. Wickramasinghe, and V. Bicknell, Mon. Not. R. Astron. Soc. **264**, 691 (1993).
26. R. Whiteburst, Mon. Not. R. Astron. Soc. **266**, 35 (1994).
27. J. Simpson, Astrophys. J. **448**, 822 (1995).
28. P. J. Armitage and M. Livio, Astrophys. J. **470**, 1024 (1996).
29. H. Yukawa, H. M. J. Boffin, and T. Matsuda, Mon. Not. R. Astron. Soc. **292**, 321 (1997).
30. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).
31. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Astron. Zh. **76**, 905 (1999) [Astron. Rep. **43**, 797 (1999)].
32. D. V. Bisikalo, P. Harmanec, A. A. Boyarchuk, *et al.*, Astron. Astrophys. **353**, 1009 (2000).
33. S. K. Chakrabarti and T. Matsuda, Astrophys. J. **390**, 639 (1992).
34. V. V. Nazarenko and L. V. Glazunova, Astron. Astrophys. (2004) (in press).
35. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **76**, 672 (1999) [Astron. Rep. **43**, 587 (1999)].
36. V. V. Nazarenko and L. V. Glazunova, Astron. Zh. **80**, 49 (2003) [Astron. Rep. **47**, 1013 (2003)].
37. V. V. Nazarenko and L. V. Glazunova, Astron. Zh. **80**, 64 (2003) [Astron. Rep. **47**, 1027 (2003)].
38. R. L. Kurucz, Astrophys. J., Suppl. Ser. **40**, 1 (1979).
39. O. M. Belotserkovskii and Yu. M. Davydov, *Large Particles in Gas Dynamics* (Nauka, Moscow, 1982) [in Russian].
40. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **72**, 190 (1995) [Astron. Rep. **39**, 167 (1995)].
41. D. V. Bisikalo, A. A. Boyarchuk, P. V. Kaigorodov, and O. A. Kuznetsov, Astron. Zh. **80**, 879 (2003) [Astron. Rep. **47**, 809 (2003)].
42. J. Binney and M. Merrifield, *Galactic Astronomy* (Princeton Univ. Press, Princeton, 1998).

Translated by K. Maslennikov

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 826–836. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 928–939.

Formation of the Accretion Disk in the SS 433 System with Explicit Radiative Cooling, Convective Heat Conduction, and Radiation Pressure

V. V. Nazarenko and L. V. Glazunova

Astronomical Observatory, National University, Shevchenko Park, Odessa, 270014 Ukraine

Received December 25, 2004; in final form, May 18, 2005

Abstract—The paper continues three-dimensional hydrodynamical computations of the formation of an accretion disk in the SS 433 system, taking into account radiative cooling explicitly, convective thermal conductivity, and radiation pressure. The computational results show that the powerful, broad flow forms an optically thick accretion disk with a gas density of 10^{12}–10^{14} cm^{-3}, a temperature of 15 000–35 000 K, a radius of about 0.3, and a height of 0.2–0.3 (in units of the component separation). Spiral shocks form in the disk, and a narrow conelike cavity (tunnel) forms at the center. In this tunnel, gas is accelerated to relativistic speeds, leaving the system in the form of narrow jets. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

We present here the second part of a study of three-dimensional hydrodynamical modeling of the formation of the accretion disk in the SS 433 binary system. In our first paper [1], we computed the formation of the accretion disk with radiative cooling specified implicitly: (1) via regularization of the temperature field, when the effects of heating in the moving medium due to numerical viscosity are restricted starting from some limited temperature, and (2) by reducing the ratio of the specific heat capacities to nearly its isothermal value, i.e., to near unity.

Our modeling of the formation of the accretion disk in SS 433 taking into account radiative cooling in various ways [1] yielded the following results. Taking into account radiative cooling via temperature regularization, we found that the flow from the point L_1, which has an initial radius of 0.2–0.3 (here and below, sizes are given in units of the distance between the components), forms an accretion disk with a radius of 0.32–0.35, which corresponds to approximately 70% of the mean radius of the accretor Roche lobe. In the orbital plane, the disk is almost circular, and its height in the z direction is roughly 0.24–0.35; i.e., it is geometrically thick and opaque. The disk's cross section with the z plane has the shape of a flattened torus. The flow forms a disk without a hot spot; i.e., the matter flows into the disk along a tangent. The interaction of gas from the disk with the outer boundary of the flow facing orbital phases 0.15–0.25 results in the formation of a high-temperature "hot line" extending from the disk to L_1. In this model, the disk does not display any definite internal structure, except for the hot line at its outer edge (between the disk and flow) and a spiral shock at the outer edge of the disk facing orbital phases 0.25–0.35. This spiral shock is quite indistinct, probably due to insufficient cooling that leads to high temperatures in the disk.

Our computations including radiative cooling in a combined approach (with both regularization of the temperature field and setting $\gamma = 1.0001$) using a high-resolution numerical grid yielded an accretion disk in SS 433 with a large number of revolutions of its gas particles. At different times, the number of revolutions varied from 70 to 150, which roughly corresponds to analytical estimates for the given accretion rate and height of the disk above the orbital plane. This model of the disk displays two lagging spiral shocks (I and II) and one leading shock (III). The disk shape is almost perfectly circular. In addition, a hot line is formed at the edge of the flow facing the disk.

After the gravitational field of the donor is removed from the area of the disk, the spiral shocks remain; however, the disk itself becomes elliptical—its radius below the line connecting the centers becomes half its radius above this line. The disk contracts by 20–30% compared to the model that includes the donor gravitation. The number of revolutions of particles in the disk decreases substantially (from 70–150 to 5), so that the disk becomes much more viscous compared to the model that includes the donor gravitation. We, therefore, conclude that (1) the gravitation of the donor affects the structure of the spiral shocks in the disk, which do not, however, disappear when this gravitation is switched off, and (2) eliminating the donor gravitation results in a substantial increase in the disk viscosity.

Since the disk temperature obtained in the study [1] is too high to be real, here we model the formation of the accretion disk in SS 433 taking into

1063-7729/05/4910-0826$26.00

account radiative cooling explicitly, using the model of optically thin cosmic plasma in the state of ionization balance [2]. In order to obtain the entire cooling pattern in the accretion disk of SS 433, we will also take into account convective heat conduction, which may be important in cooling of accretion disks.

2. MODEL FOR THE FORMATION OF THE SS 433 ACCRETION DISK WITH EXPLICIT RADIATIVE COOLING AND CONVECTIVE HEAT CONDUCTION: INITIAL AND BOUNDARY CONDITIONS, BINARY MODEL

As was noted above, in order to increase the accuracy with which radiative cooling is taken into account explicitly, we apply here a cooling model using the deexcitation function of [2]. This deexcitation function was derived from a model of optically thin cosmic plasma in a state of ionization balance.

To take into account absorption in the optically thick disk, we multiplied the deexcitation function by a factor of $e^{-\tau}$ in our computations, where τ is the optical depth in the continuum in the vicinity of the H_α line perpendicular to the orbital plane. The absorption coefficient was taken to be 10^{-17} cm^2, and the number of absorbing atoms was calculated using the Saha and Boltzmann equations. Roughly speaking, this radiative-cooling model works as follows. At high temperatures (exceeding 10 000–15 000 K), the number of absorbing atoms is low, while radiative cooling is efficient only when this number is sufficiently high (exceeding 10^8–10^{10} cm^{-3}). When the temperature goes below 10 000–15 000 K, the number of absorbing atoms rapidly increases, as does the opacity of the disk, while the efficiency of radiative cooling decreases very rapidly with increasing optical depth. Thus, the corresponding balance between the effects of numerical viscosity and radiative cooling is reached when the temperature in the envelopes is in an interval of several tens of or several thousand degrees, which is close to the real temperature in the disk and surrounding envelope.

To include cooling due to convective heat conduction, we introduce here a supplementary term in the internal-energy conservation equation for the gas (in our version of the method of large particles, the internal energy of the gas is used in the energy-conservation equation, while the standard method uses the total energy): $K_{\rm temp}\, dT^2/dz^2$, where $K_{\rm temp}$ is the coefficient of temperature conductivity and z the coordinate perpendicular to the orbital plane. This takes into account only the temperature conductivity in the z direction, but this is precisely the direction in which it is most important to estimate the temperature distribution in a geometrically thick disk. We use the coefficient of temperature conductivity $K_{\rm temp} = v_{\rm conv} r_{\rm conv}$, where $v_{\rm conv}$ is an effective convection rate and $r_{\rm conv}$ is an effective size of the convective cells.

The most difficult problem in connection with including the effect of heat conduction in the disk of SS 433 is to estimate the coefficient of temperature conductivity. We will assume that this cannot exceed 10^{-4}–10^{-5}: the maximum convection cannot exceed a quantity of the order of $v_{\rm conv} = 0.01$, and the size of the convective cells obviously cannot exceed $r_{\rm conv} = 0.01$–0.001. Here, the velocities are given in units of the orbital velocities in SS 433 (about 250 km/s). When the formation of the accretion disk is computed taking into account heat conduction, the main problem is that the time for the computations will be very long, even for the maximum temperature conductivity (for SS 433, the timescale for variations in the disk parameters will be of the order of 150–1500 orbital periods). Therefore, to enable computation of heat-conduction effects in the framework of three-dimensional hydrodynamical modeling with the currently available computers, we decided to increase the coefficient of temperature conductivity by two orders of magnitude, to the value 0.025. In this case, the time for calculating the heat-conduction effects will be approximately equal to 10–15 orbital periods. This is quite acceptable from a computational point of view; further, we will assume that all quantitative and qualitative manifestations of heat conduction (the temperature decrease in the disk and smoothing of second derivatives of the temperature with respect to the coordinates) will be the same as for the real temperature conductivity. In other words, we hope to obtain a realistic model of an accretion disk using this temperature conductivity, while reducing the necessary computation time to a reasonable level.

The binary model for the computations considered here together with its boundary and initial conditions remain the same as in [1], as do the parameters of the binary and the degree of overflow of the donor Roche lobe. We used a $150 \times 150 \times 60$ numerical grid. Each model version was calculated for 15–17 orbital periods, in order to establish a steady state in the computational domain, taking into account long-timescale convective heat conduction. The computational domain is still restricted to the vicinity of the first Lagrangian point L_1 and the accretor Roche lobe.

3. COMPUTATIONS AND RESULTS

3.1. Model for the Formation of the Accretion Disk with Explicit Radiative Cooling

As in [1], the mass-transfer rate through the vicinity of L_1 is $4 \times 10^{-5}\, M_\odot$/yr, which corresponds to an overflow of the donor Roche lobe of 0.065.

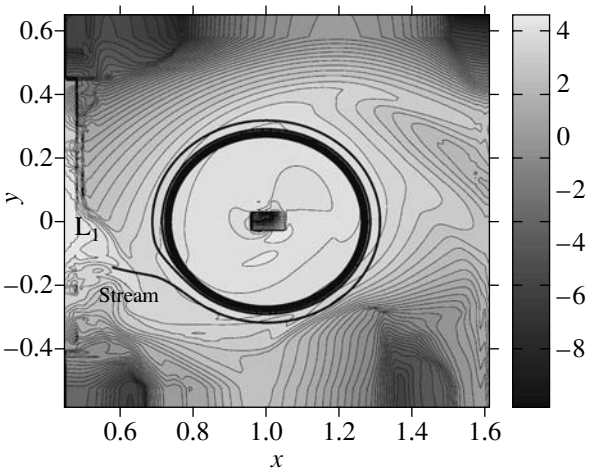

Fig. 1. Density contours and the velocity field for the cross section made by the orbital plane.

The steady state in the disk is reached approximately after a single orbital period; further, the accretion rate in the disk remains essentially constant, and its fluctuations do not exceed 0.1%. We extended our computations to more than 10 orbital periods in order to study the steady state of the disk over a longer time.

Figure 1 presents contours of constant density and a streamline in the disk and flow in the orbital plane. In the vicinity of L_1, the approximate radius of the flow is 0.10–0.15, and the disk is formed from the flow without a hot spot, but with a high-temperature "hot line." The disk displays an essentially circular shape with spiral shocks at its edge and inside the disk, close to the accretor. The size of the disk in the orbital plane is roughly 0.4, which is only slightly larger than the radius of the disk obtained in [1] (the disk occupies about 70% of the accretor Roche lobe). We, thus, conclude that the resulting disk is almost perfectly circular when radiative cooling is taken into account explicitly, as in the computations of Bisikalo *et al.* [7], with a radius that is consistent with observations.

The amount of angular momentum lost by the donor is such that it should lie on its side; i.e., the rotational axis of the donor should make an angle substantially smaller than 90° with the orbital plane. This results from the fact that, along with the z component of its angular momentum, the donor also loses its x and y components, whose absolute values are comparable to that of the z component. Thus, despite the fact that the donor had only the z component of its intrinsic angular momentum by the start of the matter outflow from L_1 (i.e., its rotational axis was perpendicular to the orbital plane), in the course of the mass transfer, the donor gains x and y intrinsic angular momentum, so that it should start tipping on its side. This result is very important for our understanding of the precession of the accretion disk, since, if the donor is on its side, it should precess on its own and drive the accretion disk into precession (forced precession [3]).

The disk displays a sharper outer boundary than in [1], apparently due to the stronger radiative cooling; in addition, the cooling is selective; i.e., it is low in denser layers of the disk close to the orbital plane and stronger in outer regions of the disk and just outside the disk. It is clear from Fig. 1 that the flow is very strongly affected by the disk, which possesses a very large specific angular moment. The reason for this may be that we have decreased the effect of numerical viscosity by including radiative cooling explicitly. Even around L_1, the disk displaces the flow below the line connecting the centers.

The density in the flow is roughly 10^{15} cm^{-3}, while the density in the disk is 10^{13}–10^{14} cm^{-3}, and in the outer envelope surrounding the disk it is 10^{9}–10^{12} cm^{-3} (meaning the parts of the outer envelope around the disk located within the computational domain).

In the given model for the accretion disk, the number of revolutions of a gas particle varies with time (from 15 to 150 revolutions; see Fig. 1). This provides evidence that the viscosity remains variable even in the steady state of the disk (for $\dot{M}_{\rm accr} = {\rm const}$); i.e., some variable processes occur in the disk, apparently generated by the rotation of the disk in the variable gravitation field of the donor. In the model, the internal structure of the disk depends strongly on the number of revolutions of the streamlines in it, i.e., on the viscosity. If the number of revolutions is small, the radial velocity is negative over the entire disk and spiral shocks are clearly visible; this is, accordingly, the case of a very viscous disk. If the number of revolutions is large, the disk possesses zones in which the radial velocity can be either negative or positive; the radial velocity in the disk alters its sign several times in the orbital plane, and the number of revolutions of the streamlines is very high precisely at these points of sign changes. The degree of concentration of the streamlines in these places is illustrated by the fact that the distance between them is only 0.00001–0.000001. Generally, in a low-viscosity state, the disk displays three zones where the radial velocity changes sign. The first is a zone around the accretor, where the radial velocity is negative (there is accretion onto a compact object); there is a zone in the middle of the disk where the radial velocity is positive (the matter flows outwards); and, above this zone, at the outer edge of the disk, there is a zone where the radial velocity is again negative. The first zone is sometimes divided into two when the viscosity is very small; in this case, a zone with positive radial velocity is situated near the inner boundary of the computational domain,

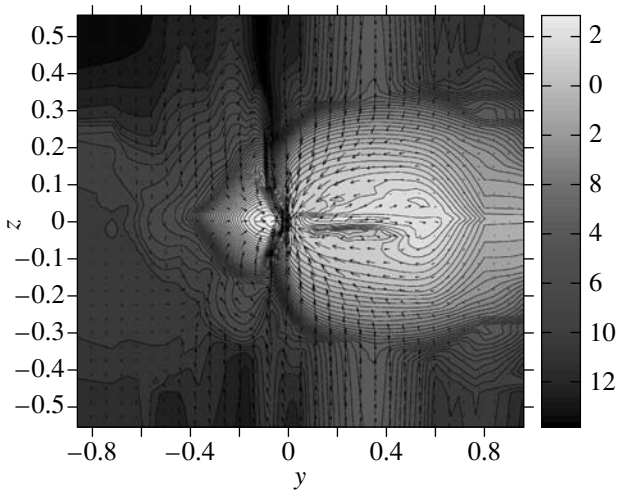

Fig. 2. Same as Fig. 1 for the cross section made by the $z-y$ plane, which contains the accretor.

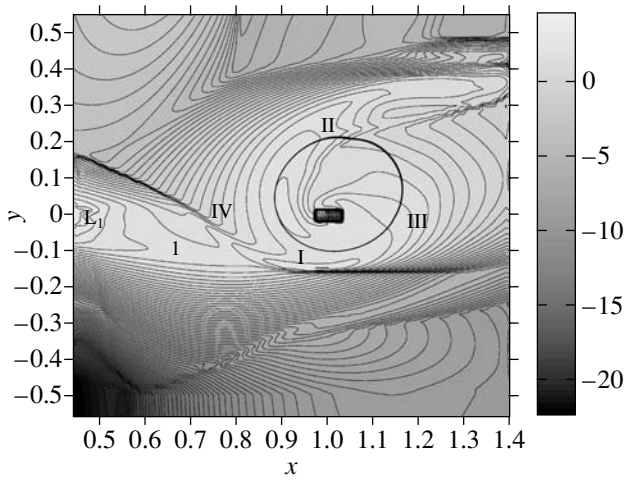

Fig. 3. Same as Fig. 1 for the disk model taking into account convective heat conduction.

approximately 0.01 from the accretor, while the sign of the radial velocity changes at a distance of 0.038. At this point, the distance between the streamlines is minimum, equal to only 0.00000001, indicating that the viscosity is very low here, probably possessing its lowest value in the disk.

Figure 2 presents density contours for the cross section of the computational domain made by the $z-y$ plane, which contains the accretor. We can see that, perpendicular to the orbital plane, the shape of the disk is toroidal, with an almost rectangular cross section. The height of the disk in the z direction remains constant and equal to roughly 0.3; only near the accretor can we see a "pit" or craterlike tunnel where the height decreases dramatically. Here, gas from the thick disk is drawn into a small volume of space in the vicinity of the accretor, forming a rarified cavity in the center of the disk. The possibility of a tunnel at the center of the disk in SS 433 has been considered in numerous studies [8]; we have confirmed its existence via direct computations of the disk structure.

Analyzing the velocity field in Fig. 2, we can see very clearly that matter moves into the tunnel; its velocity reaches a maximum at the edges of the tunnel. The disk displays a very sharp boundary in the z direction, also due to the selective model of radiative cooling used. As was noted above, the density decreases dramatically from 10^{13} to 10^{8} cm^{-3} at the disk surface. The temperature at the outer edge of the disk is roughly 35 000 K, close to the observed effective temperature of the accretion disk in SS 433 (the radiation of the disk is known to be close to that of a star with an effective temperature of about 32 000 K [4]). In the disk, the temperature increases slightly to approximately 40 000–60 000 K. The modest temperature variations indicate that the disk is almost isothermal in the z direction. However, this result cannot yet be considered final; in our opinion, a fully adequate description of the accretion disk in SS 433, and particularly calculations of the temperature structure of the disk in the z direction, require solution of the equations of radiation transfer in the disk along with the appropriate hydrodynamical equations.

3.2. Computations of the Formation of the Accretion Disk Including Convective Heat Conduction

The model computations of the formation of the accretion disk taking into account heat conduction consisted of the following stages. (1) From the start of the computations to approximately 7 orbital periods, the radiative cooling was included implicitly ($\gamma = 1.0001$) together with regularization of the temperature field ($T_{\rm lim} = 50\,000$ K). (2) Further, the computations included convective heat conduction over the entire computational domain during another 15 orbital periods, in order to establish the steady state in the disk over a sufficiently long time.

As we noted above, when heat conduction was taken into account, the formation of the disk took 10–12 orbital periods. The temperature in the disk and the second derivatives of the temperature with respect to coordinates continuously decreased. The temperature fell from a maximum of 187 000 K to 12 500 K, and the second derivative of the temperature with respect to z decreased in absolute value from 190 000 K to 3.5 K. Thus, in the steady state for the model with convective heat conduction, the temperature decreased and was, on average, 12 500 K. This is somewhat lower than the observed surface

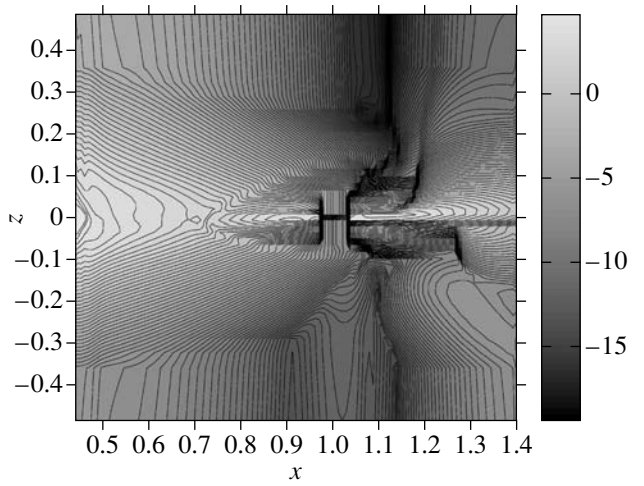

Fig. 4. Same as Fig. 3 for the cross section made by the $z-x$ plane, which contains the accretor.

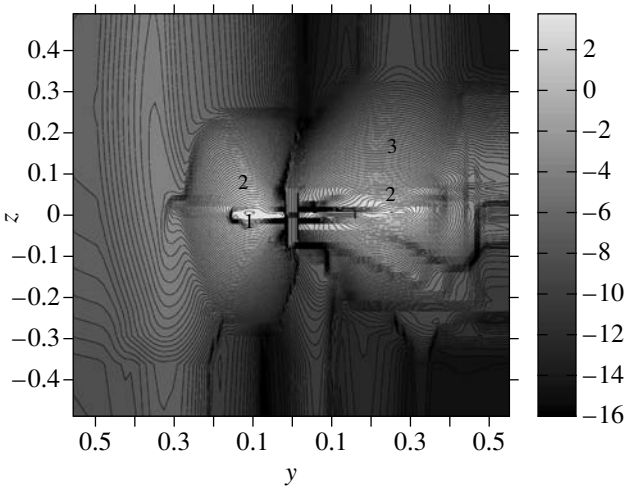

Fig. 5. Same as Fig. 3 for the cross section by the $z-y$ plane, which contains the accretor.

temperature of the disk in SS 433; however, since our calculations do not include heating of the disk by the central source, this temperature should be considered reasonably close to the observed value.

Figure 3 (orbital plane) and Figs. 4 and 5 ($z-x$ and $z-y$ planes) present the final structure of the disk in the model with heat conduction. According to Fig. 3, the radius of the disk in the orbital plane is 0.35–0.40, which coincides with our previous calculations. Generally, the structure of the disk in the orbital plane is almost the same in the model without heat conduction. Only the average temperature has decreased, as was noted above, from 187 000 K (without heat conduction and explicit radiative cooling) to 12 500 K. Even before the heat conduction was switched on, the trailing shocks I and II and leading shock III (as denoted by Bisikalo *et al.* [5, 6]) were formed in the disk. We can also see in Figs. 3–5 that the disk forms from the flow without the formation of a hot spot; instead, a hot line (IV) is formed. Initially, a "halo" is formed from the flow around the disk; matter from this halo is then gradually transferred into the disk. Essentially, in our computations, the halo does not differ from the disk material, since gas gradually moves towards the center from both places; in addition, the halo does not display density discontinuities (i.e., shocks), as occurs in the numerical model of Bisikalo *et al.* [7]. Note that the halo produced in our computations is not permanent; it only exists when the number of revolutions of a gas particle in the disk is sufficiently large (more than 15–20). Note also that, in the given model, the outer edge of the disk displays a sharp drop in density, which distinguishes the disk from the circumstellar envelope. Analyzing velocity field in the orbital plane, we can see that matter from the envelope gradually accretes onto the disk. However, the computational domain must be substantially enlarged if we wish to make final conclusions about the interaction between the disk and circumstellar envelope. Note also that the flow is not conservative in our model; analysis of the velocity field and streamlines shows that matter from the flow arrives at both the disk and the outer envelope. The circumstellar envelope in the SS 433 system is apparently formed not only due to the stellar wind from the donor but also due to matter flowing from L_1. Thus, the stellar wind from the orbital plane in SS 433 [8] is probably formed as a result of the nonconservativeness of the mass transfer from L_1.

Figures 4 and 5 present the disk structure in the $z-x$ plane containing the line connecting the centers, and the $z-y$ plane containing the accretor. Our computations have led to the formation of a geometrically thick disk with an average height in the z direction of 0.25–0.35. The structure of the disk in the $z-x$ and $z-y$ planes can be divided into three parts: (a) the core of the disk, with a density of about 10^{14} cm^{-3} (marked 1 in Figs. 4 and 5); (b) the disk itself (marked 2), with a density of $10^{12}-10^{13}$ cm^{-3}; and (c) the atmosphere (marked 3), located at heights exceeding 0.25, with a density of $10^{10}-10^{11}$ cm^{-3}. This structure explains why the disk radiates almost like a star: it has the same structure that a star should have—inner layers and an atmosphere.

It is difficult to understand the reasons for the formation of a geometrically thick disk in this system. We suggest that, since the height of the flow when it approaches the disk almost coincides with the height of the disk, which, in both cases, substantially (by a factor of two to three) exceeds the hydrostatic height of both the flow and the disk, the existence

of a geometrically thick disk in SS 433 (similar to that seen in the long-period system β Lyr [9]) is most likely due to the large initial radius of the flow from L_1. This, in turn, results from the large degree of overflow of the donor Roche lobe and the large orbital period of the system. In our opinion, it is not correct to explain the large height of the disk in SS 433 by invoking a super-Eddington luminosity for the disk (i.e., by invoking supercritical accretion), since the disk in the β Lyr system, which has similar parameters and the same mass-transfer rate through L_1, is also geometrically thick and displays no super-Eddington structure. A super-Eddington structure of the disk should also facilitate expansion in the z direction, but the contribution of this factor should not be considered to be dominant.

We compared the structure of the geometrically thick disk obtained in our calculations with the standard disk model developed by Shakura and Sunyaev [10]. As we can see from Fig. 5, the height of the disk substantially exceeds its hydrostatic height, right up to the surface of the accretor. This is what makes our disk different from the standard model, which assumes that the height of the disk is close to hydrostatic, and therefore that essentially the entire disk is illuminated by the central source. In our model with a geometrically thick disk, a large constant height is maintained in the z direction out to approximately a radius of 0.03, after which the height rapidly decreases to values below the hydrostatic height at the given point. Thus, a peculiar cavity with almost vertical walls is formed around the accretor. This is clearly visible in Fig. 5 at the center of the disk (marked 4): it is extended in the vertical direction and its walls are distinguished by powerful shocks separating the lobe from the disk itself. Recalling that the disk in SS 433 is supercritical, i.e., very intensely illuminated by the central source, we conclude that the radiation from the disk center does not illuminate the entire disk. This conclusion is based on the fact that, for a density of $10^{12}-10^{14}$ cm^{-3}, the disk should be totally opaque to the radiation from the central source, although this radiation may emerge through the cavity at the disk center, concentrated in a beam with an opening angle equal to that of the cavity. The opening angle of the cavity formed around the accretor in our computations is roughly 16°, which is close to the observed opening angle of the jets in SS 433 (about 20° [11]).

The coincidence between these angles favors our idea that jets can form in a close X-ray binary under the action of the bolometric luminosity of the disk. The key point here is that the radiation of the central source associated with the bolometric luminosity of the disk accelerates the gas around the accretor to relativistic velocities. This gas is then apparently not collimated into a beam by any specific mechanism, but is simply reflected from the walls of the narrow cavity in the disk and emerges perpendicular to the disk. Since the determination of the disk structure in the vicinity of the accretor was not the aim of this study, the formation of this cavity should be treated as a preliminary result.

On the surface of the cavity and above the cavity on the disk surface, there are sharp gradients in the density. These likely indicate the presence of shocks on these surfaces. We suggest that the emission lines observed in the spectrum of the system are formed in the vicinity of the accretor. In Figs. 4 and 5, regions with very low gas densities (below 10^5 cm^{-3}) are visible above and below the accretor. Thus, jets flowing from the craterlike central cavity should essentially expand into empty space and cannot interact with the disk, only with the wind from the disk.

Our computations show that, in the model with heat conduction, on reaching the steady state in the disk, the number of revolutions of the streamlines continuously varies erratically from 20 to 190, providing evidence for variations in the viscosity. This behavior of the disk in the steady state can explain various types of variability of the radiation from the disks in close binaries (primarily, X-ray and cataclysmic binaries). It is known that various types of variability and nonstationary behavior (outbursts, etc.) are observed in close binaries with accretion disks and compact objects, probably related to processes in their accretion disks. Studies of the nature of the disk viscosity and its variability are likely to be important for our understanding of the nature of nonstationary processes in accretion disks. We emphasize here only that the viscosity is variable in our computations, and long series of computations (on timescales of the order of 40–50 orbital periods) with high-resolution numerical grids are needed to follow variations of the internal structure of the disk and the velocity field in it in the steady-state solution in detail. We suggest that the main origin of the variability of the disk viscosity is the gravitation of the donor, since the disk rotates in a variable gravitation field.

3.3. Computations of the Formation of the Accretion Disk without the Gravitation of the Donor

In [1], we calculated the formation of the accretion disk in SS 433 without the effect of the gravitation of the donor within the disk. Initially, we included the normal gravitation of the donor and then switched it off starting from some time. The results of this approach were recalled in the Introduction. In the current study, the computations included radiative cooling explicitly and excluded the gravitation of the donor from the very start. All the parameters of the problem

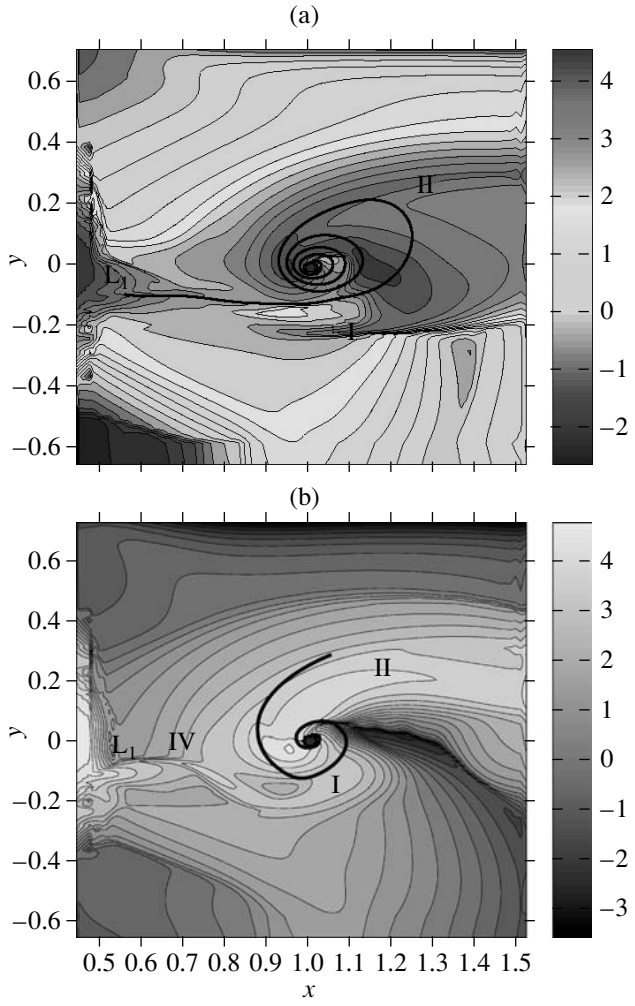

Fig. 6. Same as Fig. 1 for the disk model without the gravitation of the donor. Results for two times are presented: (a) $t = 25$, (b) $t = 45$.

remained the same, with the only difference being that we switched off the donor gravitation within a sphere with a radius of about 0.4 centered on the accretor, where it constitutes only a few percent of the accretor gravitation. Thus, our computations were carried out right from the very start without the donor gravitation. We continued the computations over seven orbital periods, in order to study the steady-state solution over a long time.

Figure 6 presents the results of these computations in the form of density contours in the orbital plane. We can see that the disk is appreciably distorted compared to the results of Subsection 3.1. The influence of the donor on the formation of the disk around the accretor is very clearly visible here. This is primarily due to the fact that the field becomes noncentral in the presence of the donor gravitation. Under the action of such a field, variations of the angular moment in the flow and disk result in dramatic differences in the disk formation compared to the case when the donor gravitation is switched off. Note also that the disk formed in this model turns out to be very viscous (the number of revolutions of the streamlines in the disk is only about two), which also illustrates the donor's influence. The variation of the disk viscosity compared to the previous results presented above results from the variation of the velocity field due to the elimination of the donor gravitation. We can see from Fig. 6 that spiral shocks still form in the disk without the donor gravitation, although their shape and structure undergo dramatic variations.

3.4. Formation of a Jet under the Action of Gravitational Pressure in SS 433

In this subsection, we will obtain a simple formula for determining the distance at which the jets originate and will estimate the velocity and density of the gas in the jets. We will suppose that the jets form under the action of collimated radiation emerging from a cone-shaped tunnel in the vicinity of the compact object, formed by a geometrically thick disk. In the standard accretion-disk model, the luminosity is given by the formula [10]

$$L(r) = \frac{GM_x M_{\rm accr}}{2r}, \qquad (1)$$

where $M_{\rm accr}$ is the disk accretion rate, M_x the mass of the compact object at the center of the disk, and r the disk radius in cylindrical coordinates (the case of disk accretion is considered here).

Suppose that, at the point where the jets originate, i.e., in the vicinity of the inner radius of the accretion disk, the disk accretion becomes spherical. This means that jets form near the spherization radius [4]. Let us also assume that the luminosity in the accretion disk is given by the same expression. In the approximation of an optically thin layer, the radiation pressure per atom is given by

$$F_{\rm rad} = \frac{\alpha L(r)}{4\pi c r^2}, \qquad (2)$$

where α is the Thomson scattering coefficient and r the disk radius in spherical coordinates. Equating the radiation pressure to the gravitation (for simplicity, we use here a Newtonian gravitation potential) yields an expression for the distance at which the jets can be formed in SS 433:

$$R_{\rm jets} = \frac{\alpha M_{\rm accr}}{8\pi c m_p}, \qquad (3)$$

where m_p is the proton mass.

To obtain the final form of this expression, we must introduce a correction for deviations from spherical

Parameters of close binaries and accretion rate in their disks from [13]

System	$P_{\rm orb}$, day	M_x, $M_\odot$	$M_{\rm opt}$, $M_\odot$	$M_{\rm accr}$, $M_\odot$/yr	$R_{\rm jets}$, cm	$V_{\rm jets}$, s	$N_{\rm jets}$, cm^{-3}	Jets
LMC X-4	1.408	1.4	16	1.0×10^{-8}	8.5×10^{6}	0.156	6.4×10^{6}	Yes
Cen X-3	2.09	1.4	18	9.0×10^{-9}	7.9×10^{6}	0.24	3.5×10^{6}	Yes
SS 433	13.086	12.0	10	1.0×10^{-7}	3.1×10^{7}	0.24	2.0×10^{11}	Yes
X1907+097	8.38	1.4	20	5.0×10^{-10}	2.0×10^{5}	1.2	8.3×10^{5}	No
LMC X-3	1.7	10	6	3.0×10^{-8}	8.5×10^{6}	0.416	4.4×10^{6}	Yes
SMC X-1	3.9	1.4	18	4.0×10^{-8}	4.1×10^{7}	0.135	7.4×10^{6}	Yes
Cyg X-1	5.6	16	33	2.5×10^{-8}	7.1×10^{6}	0.577	1.7×10^{6}	Yes
Her X-1	1.7	1.4	2.35	2.5×10^{-9}	7.1×10^{5}	0.54	3.3×10^{6}	No
X2127+119	0.713	1.4	0.9	1.0×10^{-10}	2.8×10^{4}	2.7	3.7×10^{5}	No
Cyg X-2	9.844	1.4	0.7	1.0×10^{-8}	8.5×10^{6}	0.27	3.7×10^{6}	Yes
Gen LMXB	0.2	1.4	0.5	1.0×10^{-9}	2.8×10^{5}	0.85	1.2×10^{6}	No
X1916-053	0.035	1.4	0.1	1.0×10^{-9}	2.7×10^{5}	0.83	1.2×10^{6}	No
4U 1626-67	0.029	1.4	0.03	5.0×10^{-10}	1.4×10^{5}	1.2	8.3×10^{5}	No

symmetry in the jets (which takes into account the collimation of the gas and radiation in the jets):

$$\alpha_1 = \Theta/\pi, \tag{4}$$

where Θ is the opening angle of the jet cone (if only a single jet is considered, this angle is divided by π). The final formula for the distance at which the jets are formed is

$$R_{\rm jets} = \frac{\alpha M_{\rm accr}}{8\pi c m_p \alpha_1}. \tag{5}$$

Formula (5) is obtained from (3) by dividing it by (4), since, according to our assumptions, the radiation from the disk is collimated perpendicular to the orbit plane, i.e., the radiation is concentrated in a cone. The disk luminosity increases inversely proportional to the radius; this means the luminosity decreases very rapidly outwards from the point where the jets are formed, so that the gas in the jets accelerates very rapidly to its maximum value (at infinity). Thus, the gas accelerates within a very small region within the distance $R_{\rm jets}$. This conclusion is consistent with the results of [8], where it is concluded that the gas in the SS 433 jets is accelerated to relativistic velocities within a small region of space. We can thus specify the velocity in the jet as

$$V_{\rm jets} = (2GM_x/R_{\rm jets})^2. \tag{6}$$

A similar formula is known in the theory of jets accelerated under the action of a magnetic field (see, for example, [12]); there, the distance between the central star and magnetic field is considered instead of the distance at which the jets are formed. However, MHD models and our computations differ in a very important aspect: in an MHD model, this distance is selected based on some assumption, while, in our computations, it is calculated from two parameters—the mass of the compact object and the disk accretion rate. The average density in the jets can be determined from the continuity equation, taking into account the fact that the jets form from the matter of the disk, which, in turn, originates from the stream of matter flowing from the inner Lagrangian point L_1. In other words, the average density in the jets is equal to

$$\rho_{\rm jets} = \rho_{L_1} V_{L_1}/V_{\rm jets}, \tag{7}$$

where ρ_{L_1} and V_{L_1} are the density and velocity at L_1. Thus, we have determined all the basic parameters of the jets: their density, velocity, and the distance at which they are formed. Let us use the obtained formula to consider the possibility of jet formation in other close binary systems. To this end, we will select 14 binary systems (including SS 433) from the study of Wijers and Pringle [13]. The table presents the parameters of these close binaries and the accretion rates in their disks (all data are also taken from [13]). The table also presents the parameters of jets that could be formed in these binaries. A condition for the formation of jets in a given binary is that the

distance at which they form must exceed the inner radius of the disk, which is approximately $(3-4) \times 10^6$ cm for neutron stars and $3R_g$ for black holes. In our computations, the cone opening angle at the center of the disk is assumed to be constant and equal to 20°. Using the above formula, we find that the conditions for the formation of jets are fulfilled in six of the close binaries (including SS 433).

It also follows from the table that the jet velocities for SS 433 and Cyg X-1 are very close to their observed values (0.26 for SS 433 and about 0.6 for Cyg X-1 [14] in units of the speed of light). This confirms that our estimates for the jet parameters made in this subsection are self-consistent and also in agreement with observations. According to our computations, close binaries in which jets should be formed are mainly X-ray systems with periods longer than a day and accretion rates exceeding $10^{-8} M_\odot$/yr. For short-period close binaries, the radius at which the jets are formed is smaller than the inner radius of the disk. Systems in which jets should be formed include LMC X-4, Cen X-3, SS 433, LMC X-3, SMC X-1, Cyg X-1, and Cyg X-2. The velocities of the gas motion in the jets in these systems are distributed in the interval from 0.156 to 0.577.

We can use the above formulas and the commonly accepted parameters of active galactic nuclei (AGN) to investigate the formation of jets in AGN and derive the accretion rates for which jets can be formed. For example, for a nuclear mass of $10^9 M_\odot$, jets are formed when the accretion rate exceeds $1.5 M_\odot$/yr. If with this nuclear mass the accretion rate is $10 M_\odot$/yr, the velocity in the jets formed will be approximately 0.2 of the speed of light. Thus, for reasonable parameters of AGN, the parameters of the inferred jets turn out to be close to the observed parameters, suggesting that the formulas derived here can also be used in studies of the formation of jets in AGN.

4. CONCLUSIONS

We have carried out three-dimensional hydrodynamical computations of the formation of the flow in the vicinity of the inner Lagrangian point, its motion in the Roche lobe of the accretor, and the formation of the accretion disk, taking into account radiative cooling explicitly, as well as convective heat conduction. We derived the three-dimensional structure of a steady-state accretion disk with a temperature close to the observed value. In our computations, the flow forms a disk without a hot spot, but with a hot line. This means that matter from the flow enters the disk along a tangent, and the trajectories of gas particles in the disk do not cross those in the flow. The average temperature in the disk is 25 000–35 000 K in the model with radiative cooling and approximately 12 500 K in the model with convective heat conduction.

The shape of the disk in the orbital plane is essentially circular and is slightly extended perpendicular to the line connecting the centers. The average radius of the disk is 0.35–0.40, which is approximately 70% of the average radius of the accretor Roche lobe. The temperature distribution in the orbital plane indicates that the disk displays sharp boundaries, with two shock waves being clearly visible at its outer edge: the spiral shock I, formed during the interaction of matter from the disk with the edge of the flow facing orbital phases 0.20–0.30, and the spiral wave II, which is probably formed during the interaction of gas from the outer envelope, surrounding the disk, with the edge at orbital phases 0.30–0.40. Thus, our conclusions concerning the formation of spiral shocks are consistent with those made by Bisikalo *et al.* [7], who considered the formation of spiral shocks in an accretion disk during the interaction between material in the disk and outer envelope. The analysis of the disk formation with the gravitation of the donor switched off indicates that the corresponding disk also displays a spiral structure; the shape of the disk becomes elliptical, in contrast to the shape for the model considered in Subsection 3.1, while its size remains approximately the same. The disk obtained without the donor gravitation is distinguished by its high viscosity: the number of revolutions of streamlines in the disk is only two to four. Thus, the absence of the donor's influence on the disk results in a sharper spiral structure and higher viscosity compared to the disk obtained in Subsection 3.1.

The computations of the mass transfer taking into account radiation cooling explicitly show that almost 60% of the matter flowing from L_1 flows into the disk, while roughly 40% of this matter leaves the system. When the loss of the donor's specific rotational angular momentum is calculated, it is clear that the star loses not only the z component but also the x and y components of this vector. This indicates that the donor should lie on its side, which is important in connection with explaining the precession of the accretion disk. Indeed, in the most popular model for the disk precession, the disk follows the precession of the donor. As long as the donor lies on its side, it should itself precess, giving rise to precession of the accretion disk. Thus, our computations suggest that precession of the accretion disk is simply a consequence of the matter flow through the vicinity of L_1.

The structure of the disk in the z direction indicates that the disk shape in the $z-x$ and $z-y$ planes is close to a torus with a transverse cross section with a complex shape and with a height of 0.24–0.34,

which exceeds the hydrostatic height by a factor of two to three (only in the vicinity of the accretor is a depression in the disk seen, where the height of the disk is close to the hydrostatic value of about 0.05–0.08); i.e., our computations imply the formation of a geometrically thick disk. A rarified conelike cavity is formed near the polar regions, at the center of the disk—a tunnel, which is often considered in connection with the formation of the SS 433 jets [8]. The cone opening angle in our computations is roughly 16°, which coincides with the opening angle of the SS 433 jets (about 20°). This indicates good agreement between our disk model and the observations [8].

The reason for the formation of this tunnel in the disk center is very simple: the matter from a geometrically thick disk should be accreted into a small area at the center of the disk, which results in the formation of a cavity, along whose surface gas moves towards the accretor.

Since the initial height of the flow is roughly 0.2–0.3, we connect the formation of a thick, opaque disk in SS 433 with the large initial radius of the flow, which, in turn, results from the high degree of overflow of the donor Roche lobe and the long period of this binary.

Our study shows that the gas near the inner edge of the disk can be accelerated to relativistic velocities under the action of the bolometric luminosity of the disk (we derived a formula for determining the radius at which relativistic jets can be formed based on the standard Shakura–Sunyaev model for the disk luminosity). We assume that the focusing of the jets in SS 433, as well as in other close binaries and in AGN, where thick disks may be formed, is not driven by any specific mechanism; gas accelerated to relativistic velocities simply leaves the system through a narrow cavity at the center of the disk.

The temperature and density distributions in these planes also indicate that the disk displays a sharp surface in the z direction (the density in the disk is 10^{13}–10^{14} cm^{-3} at its center, 10^{12}–10^{13} cm^{-3} in the disk itself, and from 10^{12} to 10^{9} cm^{-3} at its edge, i.e., in its atmosphere). Thus, according to our models, the disk in SS 433 displays a complex structure in the z direction, similar to that of an ordinary star (with inner layers and an atmosphere). This is consistent with observational data, which indicate that the disk in the SS 433 system radiates like a star with a surface temperature of 32 000 K. The temperature in the disk obtained in our computations with radiative cooling is about 25 000–38 000 K in the inner parts of the disk and roughly 12 000–19 000 K in its outer parts, in good agreement with the observations.

The most important conclusions from our study are the following.

(1) The radius of the flow formed from the atmosphere of the donor in the vicinity of the first Lagrangian point L_1 is 0.15–0.20 for a mass-transfer rate through this point of the order of $3 \times 10^{-5}\ M_\odot$/yr. A geometrically thick disk with a height of 0.2–0.3 forms from this flow. This can explain the formation of the thick disk in SS 433 by purely astrophysical (the large mass of the donor and large degree of overflow of its Roche lobe, the long orbital period of the system) rather than physical reasons (for example, the presence of strong viscosity).

(2) The mass transfer through L_1 turns out to be nonconservative—roughly 60% of this matter flows into the disk, while 40% leaves the system.

(3) The flow forms an almost circular disk without a hot spot, but with a hot line. The disk formed from the flow has a radius of roughly 0.3 with a matter density of about 10^{12}–10^{14} cm^{-3} and a temperature of 15 000–35 000 K. The disk height in the z direction is substantially larger than the hydrostatic height and is equal to approximately 0.3.

(4) Spiral shocks are formed in our disk model, probably during the interaction between the disk and flow, as well as between the disk and envelope.

(5) The gas in the inner parts of the disk can be accelerated to relativistic velocities under the action of the bolometric luminosity of the disk.

(6) A rarified conelike cavity (tunnel) is formed at the center of the disk; gas from the thick disk is accreted into a small region at the center of the disk along the walls of this cavity.

(7) Due to the existence of this cavity, gas accelerated to relativistic velocities simply leaves the disk through the cone-shaped cavity, instead of being focused into narrow jets by some specific mechanism.

(8) The distances at which jets can be formed in the disks of X-ray binaries, estimated using formula (5), indicate that jets can be formed in the disks of seven close binaries, including Cyg X-1 and SS 433, and that the gas in these jets can move outward with velocities from 0.155 to 0.599 of the speed of light.

(9) The computations show that a high-temperature region (roughly 90 000–110 000 K) is formed above the accretor. Such regions also originate in computations of mass transfer in close binaries with ordinary components (β Lyrae, RZ Sct, V367 Cyg, U Cep systems). We relate the existence of these regions with the formation of emission lines in the spectra of some close binaries, in particular, SS 433.

REFERENCES

1. V. V. Nazarenko, L. V. Glazunova, and S. V. Nazarenko, Astron. Zh. (2005) (in press).
2. D. P. Cox and E. Daltabuilt, Astrophys. J. **167**, 113 (1971).
3. A. M. Cherepashchuk, Space Sci. Rev. **102**, 23 (2002).
4. M. A. Abramowitz, N. E. White, and L. G. Filipov, *The Physics of Compact Object* (Pergamon, Oxford, 1989).
5. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).
6. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Astron. Zh. **76**, 905 (1999) [Astron. Rep. **43**, 797 (1999)].
7. D. V. Bisikalo, A. A. Boyarchuk, P. V. Kaĭgorodov, and O. A. Kuznetsov, Astron. Zh. **80**, 879 (2003) [Astron. Rep. **47**, 809 (2003)].
8. S. N. Fabrika, Astrophys. Space Phys. **12**, 1 (2004).
9. V. V. Nazarenko and L. V. Glazunova, Astron. Zh. **80**, 49 (2003) [Astron. Rep. **47**, 1013 (2003)].
10. N. T. Shakura and R. A. Sunyaev, Astron. Astrophys. **24**, 337 (1973).
11. W. K. Rose, Mon. Not. R. Astron. Soc. **276**, 1191 (1995).
12. T. Kudoh and K. Shibata, Astrophys. J. **474**, 362 (1997).
13. R. A. M. J. Wijers and J. E. Pringle, Mon. Not. R. Astron. Soc. **308**, 207 (1999).
14. A. M. Stirling, R. E. Spencer, C. J. de la Force, *et al.*, Mon. Not. R. Astron. Soc. **327**, 1273 (2001).

Translated by K. Maslennikov

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 837–846. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 940–949.

The Spectrum of Solar Cosmic Rays: Data of Observations and Numerical Simulation

Yu. V. Balabin[1], É. V. Vashenyuk[1], O. V. Mingalev[1], A. I. Podgorny[2], and I. M. Podgorny[3]

[1] *Polar Geophysical Institute, Apatity, Russia*

[2] *Lebedev Physical Institute, Moscow, Russia*

[3] *Institute of Astronomy, Moscow, Russia*

Received October 30, 2004; in final form, February 17, 2005

Abstract—Analysis of the relativistic proton spectra of solar flares occurring in the 23rd solar activity cycle derived from data of a worldwide neutron monitor network and numerical modeling both provide evidence for the acceleration of charged particles by an electric field that arises in coronal current sheets during reconnection. The method used to obtain the spectra is based on simulating the response of a neutron monitor to an anisotropic flux of relativistic solar protons with specified parameters and determining the characteristics of the primary relativistic solar protons by fitting model responses to the observations. Studies of the dynamics of the energy spectra distinguish two populations of relativistic protons in solar cosmic-ray events: the so-called fast component, which arrives at the flux front of the solar cosmic rays, followed by the delayed slow component. The fast component is characterized by strong anisotropy and an exponential energy spectrum, in agreement with the spectrum yielded by mathematical modeling of particle acceleration by an electric field directed along the X line of the magnetic field. The slow component, whose propagation is probably diffusive, has a power-law spectrum. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

A solar flare is made up of a collection of complex processes arising on the Sun. It has now been firmly established that the primary flare energy release occurs in the corona, with the duration of the explosive phase of a flare that releases up to 10^{32} erg being usually of the order of 100 s. The main manifestations of such flares are

—the appearance on the solar surface of "ribbons" glowing at visible wavelengths in an active region, with the distance between the ribbons of two-ribbon flares growing with time;

—the appearance in an active region of centers of soft and hard X-ray radiation associated with the fast electrons precipitating onto the chromosphere;

—the ejection of coronal material into interplanetary space and the associated formation of a shock wave;

—the generation of fluxes of protons that are accelerated to energies exceeding 1 GeV;

—the generation of γ radiation associated with nuclear reactions occurring during the flare.

The 2.23 MeV line due to the radiative capture of a neutron by a proton is clearly observed in the γ-ray spectrum.

The flux of high-energy protons arrives at the Earth along lines of the spiral interplanetary magnetic field, where they can then penetrate the magnetosphere of the Earth. Protons with sufficiently high energies, exceeding the so-called geomagnetic cutoff threshold, reach the Earth's atmosphere, giving rise to fluxes of neutrons formed in nuclear reactions, which can be detected by neutron monitors at the Earth's surface. The network of neutron monitors distributed around the Earth can be considered as a single multidirectional instrument for the measurement of the spectra, anisotropies, and pitch angles of fluxes of primary relativistic solar cosmic rays arriving from beyond the Earth's magnetosphere [1, 2].

In the current study, we derived the spectra of relativistic solar protons (RSPs) by applying fitting methods to data from a worldwide network of neutron monitors. The analysis of these data also included computation of the particle trajectories based on modern models for the Earth's magnetosphere.

The temporal characteristics of fluxes of electromagnetic radiation and protons detected at the Earth's orbit differ appreciably. The duration of the γ-ray pulse coincides with the duration of the explosive phase of the flare, while the duration of the flux of protons is measured in hours, with the growth front of the flux sometimes being very sharp. Some of the protons move into interplanetary space, arriving at the Earth in the flux front, and some give rise to nuclear reactions on the Sun, accompanied by the

1063-7729/05/4910-0837$26.00

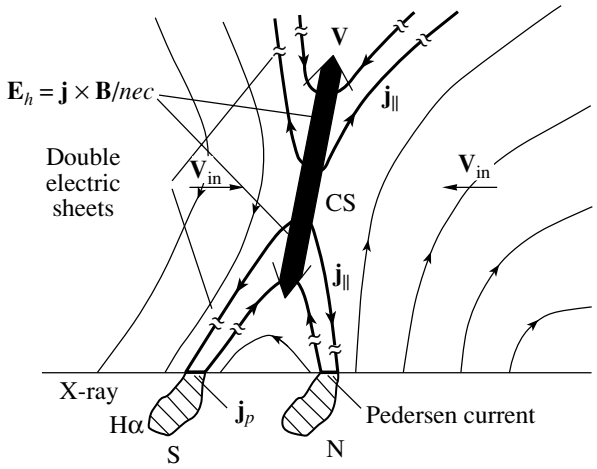

Fig. 1. Electrodynamical model for a solar flare. Lines of the magnetic field (thin) and longitudinal currents are shown, together with velocity vectors for the plasma. The electric field $\mathbf{E} = -\mathbf{V} \times \mathbf{B}/c$ is perpendicular to the plane of the figure.

emission of γ rays. The origin of the delayed protons (the so-called slow component) may not be related directly to processes occurring during the flare. For example, they could be accelerated in a shock wave created by a supersonic coronal ejection. In this case, the energy spectra of the protons arriving in the flux front and the flux tail should be different. It is also possible that the slow component is made up of particles that were captured in the magnetic field of the active region, whose spectrum is formed during interactions with turbulence.

The appearance of a collection of flare phenomena can be explained by electrodynamical models for the flare [3–5], which assume the accumulation of energy in the corona above an active region in the magnetic field of a current sheet. During the evolution of a current sheet in an unstable state and its subsequent disruption, the energy of the magnetic field transforms into kinetic and thermal energy. This flare model is illustrated in Fig. 1. The thin curves depict the magnetic-field lines; a current sheet is contained between the oppositely directed field lines. The larger arrows show the plasma velocity. The plasma flows into the current sheet from both sides together with the magnetic-field lines and is accelerated along the sheet by the $\mathbf{j} \times \mathbf{B}/c$ force after magnetic reconnection in the vicinity of the X line of the field (j is the current, B the magnetic field, and c the speed of light). The plasma accelerated upward creates a coronal mass ejection. Together with contracted lines after reconnection, the downward flow of plasma heated during the reconnection forms hot postflare loops. The Hall electric field directed along the sheet, $\mathbf{E} = \mathbf{j} \times \mathbf{B}/nec$ (n is the particle number density, and e the charge per particle), gives rise to a system of longitudinal currents (parallel to the magnetic field lines) that close in the chromosphere. The electrons that are accelerated in these currents precipitate onto the solar surface, leading to the visible and X-ray emission of the flare ribbons. Protons can be accelerated to high energies in the flare along the X line of the magnetic field, perpendicular to the plane of the figure. This acceleration occurs in the $\mathbf{E} = -\mathbf{V} \times \mathbf{B}/c$ electric field, which grows in the case of rapid reconnection during the disruption of the current sheet. Here, $\mathbf{V}$ is the velocity with which plasma flows into the sheet and $\mathbf{B}$ is the magnetic field of the current sheet. In general, the X line is not a neutral line. The magnetic field directed along this line should play a stabilizing role during the acceleration processes. Estimates show [4] that, in principle, this mechanism should be able to accelerate protons to energies substantially exceeding those typically registered during flares.

The most trustworthy data on the development of a flare were obtained from RHESSI measurements for the flare of July 23, 2002 [6]. The 40 keV X-ray emission appeared nine minutes before the explosive phase of the flare. The X-ray spectra indicate that this was not thermal emission, but was due to electrons accelerated in the corona, which partially precipitated onto the chromosphere, as is expected in the electrodynamical model of a flare [5]. Further, in the explosive phase, there is a sharp growth in the X-ray emission at energies exceeding 100 keV and the γ-ray emission. Simultaneously, thermal X-ray emission from a coronal source with the temperature of $\sim 4 \times 10^7$ K was detected, testifying to the heating of plasma in a current sheet during reconnection. The motion of ribbons of hard X-ray emission was observed at this same time, likewise providing evidence for reconnection and the precipitation of electrons onto the chromosphere in the feet of loops [7]. Measurements of γ-ray lines and the γ-ray continuum indicate the acceleration of ions and electrons to energies of no lower than tens of MeV in the explosive phase. The γ-ray emission is associated with accelerated protons arriving at the solar surface. The acceleration of protons and emission of γ rays occurs simultaneously in the impulsive phase of the flare. We note in passing that the appearance of a coronal X-ray source could precede the precipitation of electrons onto the photosphere, as was the case for the flare of January 24, 1992 [8]. Apparently, the energy that is released during reconnection can be distributed between the accelerated electrons and the hot plasma in the coronal source in various ways, depending on the conditions during the development of the flare.

Thus far, theoretical analyses of the acceleration of particles during the disruption of a current sheet have

been carried out under very simplified conditions, leading to sometimes contradictory results [9–12]. The acceleration has been assumed to occur in an externally applied electric field, with the current sheet replaced by a stationary existing magnetic-field configuration with a neutral X line. In [6], the energy spectrum of the protons was determined from the motion of particles in the vicinity of the X line. A particle moving precisely along the X line is accelerated by the full applied potential difference, while a particle that initially moves at some distance δ from the X line is deflected by the magnetic field and gradually leaves the region of nonadiabatic motions without acquiring its maximum possible energy. The larger is δ, the smaller the potential difference passed through by the particle. This yields a spectrum of the form $dN/dW \sim \exp(-W/EeL)$ (W is energy and L is distance). However, according to [10], the spectrum for particles accelerated in the region beyond the nonadiabatic region could be a power law. A power-law spectrum can be formed for particles with modest energies. Numerical simulation was carried out in [11] based on the analysis of [9], leading to the conclusion that the spectrum should have a power-law character, as is typical for the spectrum of cosmic rays. In addition, the computations of [12], carried out in a similar fashion, indicate an exponential spectrum for the accelerated protons, in agreement with the experimental data of [2, 13].

A possible mechanism for the generation of the slow component is stochastic acceleration by plasma turbulence [13]. The goal of the current study is to determine the spectra of solar cosmic rays at various times based on data obtained on a worldwide network of neutron monitors and to compare these spectra with computed spectra for protons accelerated in the vicinity of an X-type magnetic-field line in a current sheet, when the electric field is determined by magnetic reconnection.

Spectra of RSPs in a series of events of the current (23rd) solar-activity cycle are analyzed using a method based on computing the trajectories of protons of various energies that give rise to the detected neutrons [1, 2]. We simulated the response of the worldwide network of neutron monitors to an anisotropic flux of RSPs with specified parameters and compared the results to observations. This procedure includes computing the trajectories of the protons in a model geomagnetic field. In each of three events analyzed (July 14, 2000, October 28, 2003, and November 2, 2003), the parameters of the RPSs were determined at various times, enabling us to derive data on the dynamics of the fluxes and spectra of these protons beyond the magnetosphere. In each case, two different populations of RSPs can be distinguished—prompt and slow—with the former being observed in the initial phase of an event. The fast component reaches the Earth's orbit over the flight time of the protons traveling along the spiral lines of the interplanetary magnetic field. This means that the spectrum of the prompt component is not distorted by diffusion due to collisions between the protons and magnetic inhomogeneities and does not contain particles captured by the strong magnetic fields of the active region. Thus, the spectrum of the prompt component carries information about the particle-acceleration mechanism in the solar flare. The prompt component is characterized by an impulsive temporal profile that reflects the temporal characteristics of the flare, as well as pronounced anisotropy.

The delayed slow component arrives at the Earth later, and begins to dominate in the flux of RSPs 0.5–1 h after the onset of the event. The mechanism determining the spectrum of the slow component may be acceleration in a shock or the escape of fast particles from magnetic loops that are captured by the magnetic field during the flare. We also cannot exclude the possibility that the slow component spectrum is influenced to some extent by the diffusion of particles due to collisions with magnetic inhomogeneities in the interplanetary medium. A firm answer about the nature of the associated particle acceleration can only be provided by comparisons of computed spectra for various acceleration models with observational data. In the current study, we have carried out numerical simulation of spectra obtained for the case when particles are accelerated by an electric field directed along a neutral line of the magnetic field and compared the resulting spectra with neutron-monitoring measurements. Neutron monitors currently provide possibly the only means of measuring the spectra and other characteristics of relativistic solar cosmic rays, which are not accessible to direct measurements on spacecraft or balloons due to the small effective areas of the detectors that can be used. Our results indicate that the computed spectra are in agreement with the spectra of the RSP prompt component derived from neutron-monitor data.

2. METHOD FOR DETERMINING THE SPECTRA OF THE SOLAR PROTONS

The response of a neutron monitor to an anisotropic flux of RSPs beyond the Earth's magnetosphere is a complex function of the characteristics of the flux itself, the conditions for the propagation of the particles in the Earth's magnetosphere, and the conditions for the passage of neutrons generated in the upper layers of the atmosphere by the primary protons through the atmosphere. We derived the energy spectra of the RSPs together with other characteristics of the flux of these particles in the interplanetary space

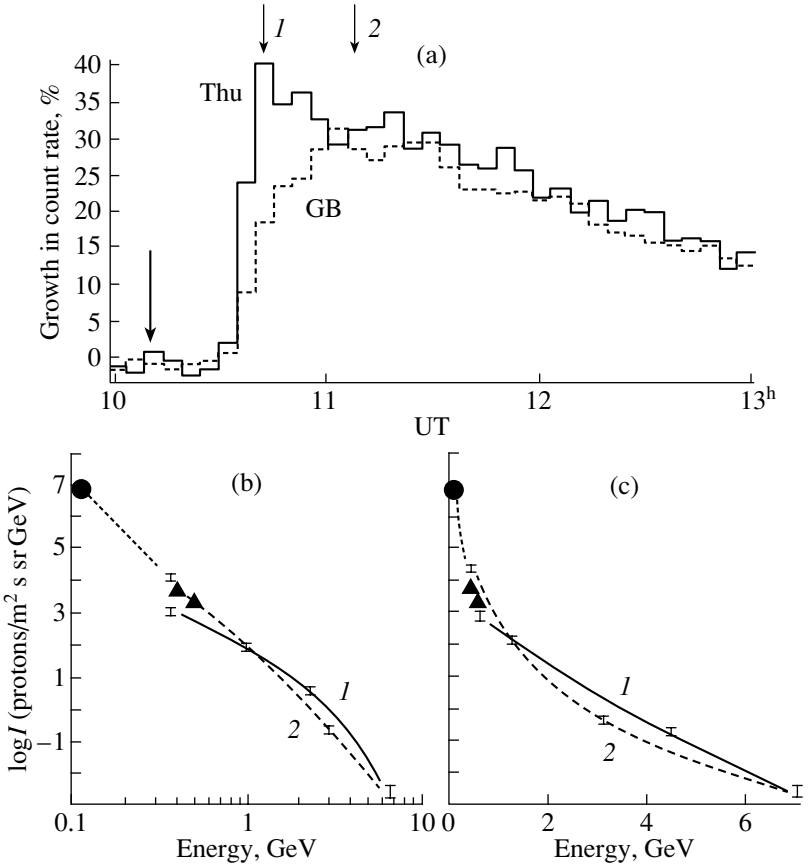

Fig. 2. (a) Profiles of the growth of the neutron flux at the neutron monitoring states at Thule (Thu) in Greenland and Goose Bay (GB) in Canada during the event of July 14, 2000. The bold arrow shows the supposed time when the protons were generated. The times when the spectra were calculated are also indicated: (*1*) for the prompt component and (*2*) for the slow (delayed) component. (b, c) The spectra on log–log and semilogarithmic scales. The errors in the measurements are indicated. The direct measurements of the solar protons are shown by crosses (balloon probes) and circles (the GOES-11 spacecraft).

(their anisotropy, the pitch angle distribution in the interplanetary magnetic field near the boundary of the magnetosphere) by fitting computed responses for a worldwide network of neutron monitors to the observations [2]. In this case, we must know the asymptotic directions (the direction of arrival at the boundary of the magnetosphere) of protons with a specified energy (hardness) that contribute to the signal for a given neutron monitor. For each station, we calculated the asymptotic reception cones formed by the asymptotic arrival directions of protons in steps of 0.001 GV in the hardness range from the atmospheric cutoff $pc \sim 1$ GV (proton energies of 430 MeV) to $pc = 20$ GV (proton energies of 19 GeV), which is the limiting hardness in the spectrum of solar cosmic rays.

The computations were carried out by integrating the equation of motion of a particle with negative charge and the mass of the proton with a specified hardness emitted vertically upward from a height of 20 km above a given station. The height of 20 km was chosen since it is the mean altitude at which the secondary neutrons that contribute to the neutron-monitor signal are formed. The equations of motion were integrated using a fourth- or fifth-order Runge–Kutta method. The computations used the Tsiganenko 2002 magnetospheric model. Using the calculated asymptotic directions, we computed the response of a given neutron monitor to an anisotropic flux of RSPs with specified parameters. By comparing the computed responses with actual measurements via a least-squares fit, we found a set of parameters for the flux of RSPs that yielded the best agreement with the observations for the entire network of neutron-monitoring stations. The number of stations used for the analysis was from 27 to 30; this also corresponds to the number of conditional equations used in the fitting.

3. DYNAMICS OF THE SPECTRA OF SOLAR FLARES BASED ON NEUTRON-MONITOR DATA

3.1. Event of July 14, 2000

Figure 2 presents temporal profiles showing the growth of the signals at the two terrestrial neutron-

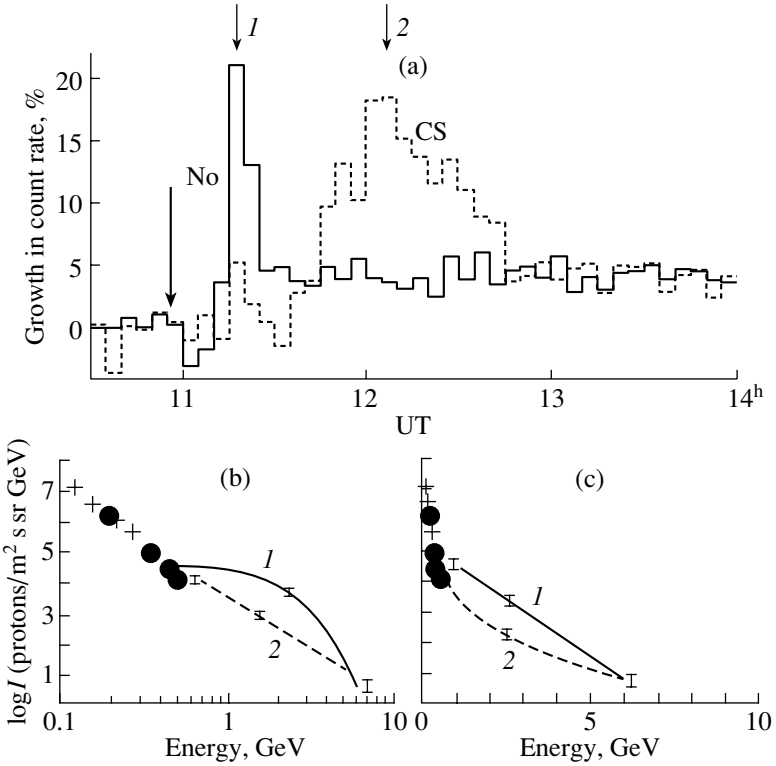

Fig. 3. Same as Fig. 2 for the event of October 28, 2003 and the neutron-monitoring stations at Norilsk (No) and Cape Schmidt (CS). We can see that the Norilsk station detected predominantly the prompt component, while the Cape Schmidt station detected primarily the slow component.

monitoring stations at Thule (Thu) in Greenland and Goose Bay (GB) in Canada during the RSP event of July 14, 2000 (the "Bastille Day flare"). These profiles show the presence of two populations of RSPs: prompt and delayed (slow). The smooth growth in the count rate at Goose Bay indicates that this station detected primarily the slow component, while a sharp initial peak corresponding to the prompt component has been recorded at the Thule station. The arrows labeled *1* and *2* in Fig. 2a mark the maxima in the count rates recorded at these stations.

Figures 2b and 2c present the energy spectra of the RSPs for these times (shown by the arrows in Fig. 2a). The spectra were derived by fitting the data for 27 neutron-monitoring stations [2] and are presented on log–log and semilog scales. These data show that spectrum *1*, which was obtained for the initial phase of the event (the fast component), is described well by a linear dependence on a semilog scale (Fig. 2c); i.e., this relation is exponential. Spectrum *2*, which was taken 30 min later, has a different character and can be described by a linear dependence on a log–log scale; i.e., it is a power-law dependence with a high index, $\gamma \sim -5$.

3.2. Event of October 28, 2003

The presence of two components in the RSPs is demonstrated in the event of October 28, 2003, by the data for the two high-latitude neutron-monitoring stations Norilsk (No) and Cape Schmidt (CS). As we can see from Fig. 3a, a sharp maximum in the count rate (arrow *1*) due to the arrival of the fast component is clearly observed at the Norilsk station; at Cape Schmidt, a modest peak due to the arrival of the fast component is first detected (arrow *1*), while a flux of neutrons with a shallow front associated with the slow component (arrow *2*) is detected over a prolonged time beginning after a delay of ~30 min. The strong difference in the time dependences is due to the high anisotropy in the flux of relativistic solar cosmic rays. The amount of signal growth at different stations with the same geomagnetic threshold depends on the orientation of their asymptotic cones relative to the interplanetary magnetic field and the shape of the pitch-angle distribution. This last quantity evolved rapidly during the event of October 28, 2003: a strong variation in the pitch-angle distribution of the particles from the initial phase of the event (the fast component) to the time of the maximum of the slow component.

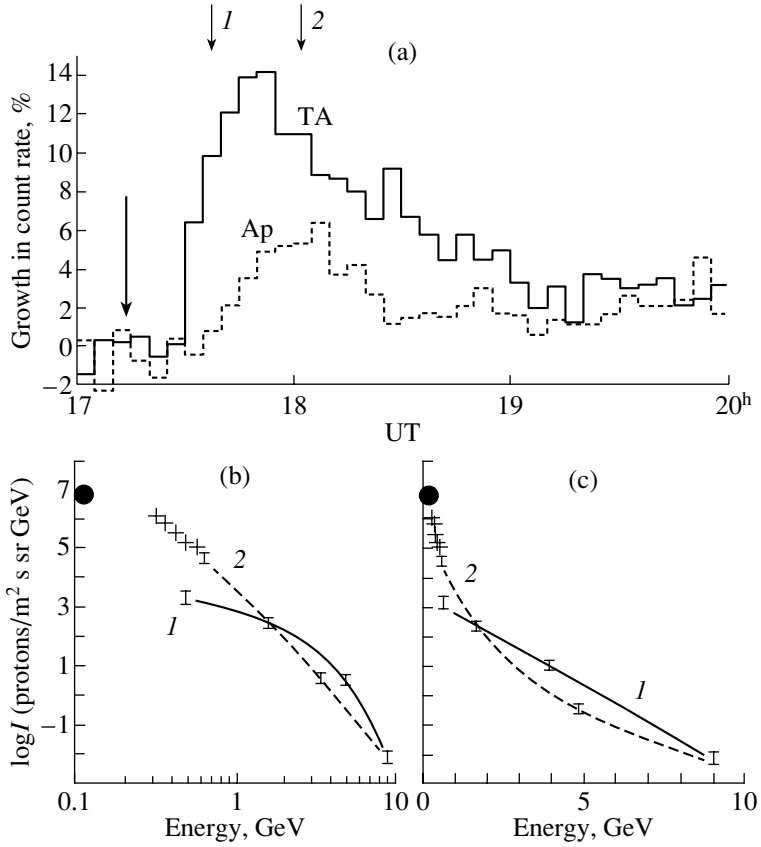

Fig. 4. Same as Fig. 2 for the event of November 2, 2003, and the neutron-monitoring stations at Apatity (Ap) and Terra Adelaide (TA) in Antarctica. Only the slow component was detected at Apatity, while a superposition of the prompt and slow components was detected at Terra Adelai.

Figures 3b and 3c show the solar-proton spectra on log–log and semilog scales derived by fitting the data for 30 neutron-monitoring stations at the times *1* and *2* marked by the arrows in Fig. 3a. The spectrum of the RSPs for the fast component obtained at 11:20 UT (curve *1*) differs strongly from the spectrum obtained at 12:00 UT, when the flux of solar rays has been dominated by the slow component, giving rise to the appearance of a delayed maximum. The linear character of the spectrum on the semilog scale for the fast component shows that, as in the previous case, the spectrum is exponential, while we see a linear energy dependence for the slow component on the log–log scale; i.e., the spectrum of the delayed protons is a power law. The power-law index ($\gamma \sim -4$) is smaller than that for the Bastille Day flare. The crosses and circles represent direct measurements of moderate-energy solar protons made on board the GOES-10 spacecraft and using balloon probes during this event (http://goes.ngdc.noaa.gov/data/avg/2004/), which correspond to times $t >$ 12:00 UT. The spectra derived from these measurements at energies of 100–300 MeV are approximated well by the spectra derived from the neutron-monitor data. Note that the GOES-10 measurements can only be used to analyze a narrow interval of energies, in which it is difficult to unambiguously distinguish between a power law and exponential spectrum.

3.3. Event of November 2, 2003

Figure 4a shows growth profiles obtained for the Apatity (Ap) station and the Terra Adelie (TA) station in the Antarctic. Only the slow component of the RSPs is present in the Apatity profile. The presence of a sharp front in the growth of the neutron flux at the Terra Adelaide station is associated with the arrival of the prompt component, while a slowly decreasing flux of neutrons (the slow component) is detected after the maximum (17:45 UT). The dynamics of the spectra for this event (Figs. 4b, 4c) are similar to those for the event of October 28, 2003. The spectrum of the prompt component has a linear character on the semilog scale (curve *1*)—i.e., it is exponential—while the spectrum for the slow component is described well by a power law. Spectrum *1*, which corresponds to the fast component, is harder.

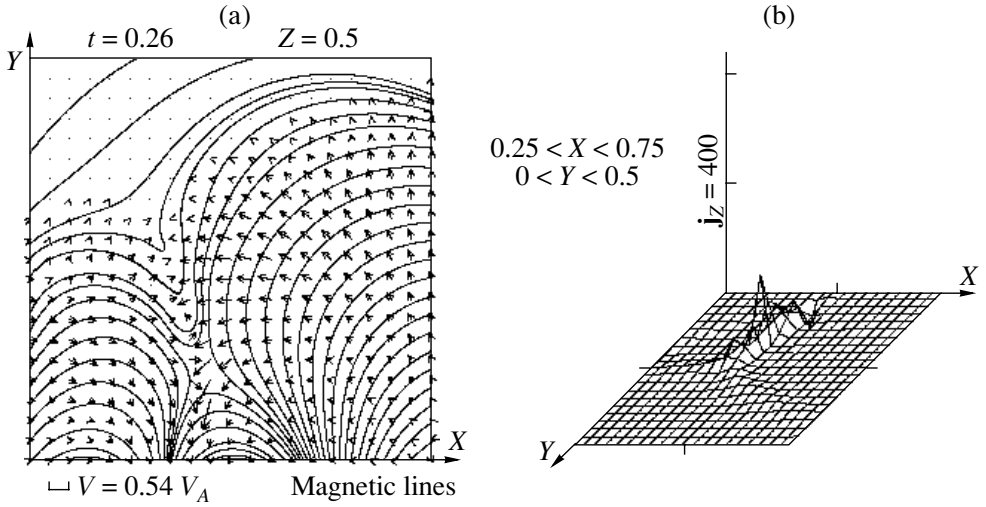

Fig. 5. Results of numerical MHD simulation of the Bastille Day solar flare: (a) magnetic-field lines and velocity vectors; (b) distribution of current density in the current sheet at the onset of the flare.

4. NUMERICAL SIMULATION OF THE SPECTRA OF PROTONS ACCELERATED DURING THE FLARE

The maximum possible energy of solar cosmic rays accelerated by the electric field generated during the disruption of a current sheet is determined by the total potential difference applied to the singular line of the magnetic field. However, in the case of an extended singular line, this potential difference can be passed through by only a small number of particles, so that the real maximum energy that is detected will be appreciably smaller. We estimate the maximum possible energy using the results of three-dimensional numerical MHD simulations of the Bastille Day flare [14, 15]. Figure 5 presents the results of such simulations of the field of this flare (July 14, 2000). The figure shows the velocity distribution and magnetic-field lines of the developing flare (Fig. 5a), as well as the distribution of the current in the reconnecting current sheet (Fig. 5b). In this calculation we assume: the flow velocity in the sheet is $\sim 0.1\, V_A \sim 10^9$ cm/s, the field in the current sheet is ~300 G, and the maximum electrical field is $-\mathbf{V}\times\mathbf{B}/c \sim 3\times 10^3$ V/cm, and if the length of the singular line is $L \sim 10^9$ cm, the maximum possible energy is 3×10^{12} eV. Such energies can be attained only by particles that have traversed the full potential difference, i.e., that have not left the acceleration region. When L is large, the number of such particles is negligibly small. The computation of the proton spectrum during the Bastille Day flare is currently being prepared. As a first step in this direction, we have numerically calculated the spectra for the case of acceleration along the singular line, with the potential difference being applied along this line.

The computation of the proton acceleration was carried out in the closed parallelepiped

$$\Pi = \{\mathbf{x} : |x_i| \le L_\perp, \quad i = 1,2, \quad 0 \le x_3 \le L_{||}\}, \tag{1}$$

with constant electric and magnetic fields specified in the form

$$\mathbf{E} = E_0\mathbf{e}_3 = -\nabla\varphi(\mathbf{x}), \quad \varphi(\mathbf{x}) = -E_0 x_3, \tag{2}$$
$$\mathbf{B} = (B_0/L_\perp)\,(x_2\mathbf{e}_1 + x_1\mathbf{e}_2),$$

where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the Cartesian basis vectors in R^3 space. The magnetic field specified in this way contains the zero line $\{x_1 = 0, x_2 = 0\}$. This model field was used to simulate the particle acceleration in [11, 12].

The following model was used for the accelerated-particle spectrum for the numerical modeling. The parallelepiped Π was divided along the x_3 axis into $\nu_{||}$ layers $\Pi(k)$ with thickness L_g:

$$\Pi = \bigcup_{k=1}^{\nu_{||}} \Pi(k), \quad \Pi(k) = \{\mathbf{x} : |x_i| \le L_\perp, \tag{3}$$
$$i = 1,2, \quad (k-1)L_g \le x_3 < kL_g\}.$$

In turn, each layer was separated into $\nu_\perp^2$ cubes with sides of length L_g. (Thus, the lengths $L_\perp$ and $L_{||}$ should be multiples of L_g, namely, $L_\perp/L_g = \nu_\perp$, $L_{||}/L_g = \nu_{||}$.) In each such cube of the first layer $\Pi(1)$, we specified $2^{12} = 4096$ protons and numerically generated the three-dimensional spatially uniform Maxwellian velocity distribution

$$f(\mathbf{v}) = n\left(V_T\sqrt{2\pi}\right)^{-3}\exp\left(-\mathbf{v}^2/\left(2V_T^2\right)\right) \tag{4}$$

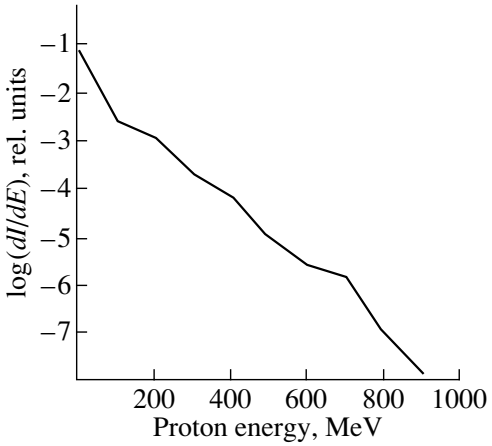

Fig. 6. Computed spectrum of protons accelerated by an electric field lying along an X-type neutral line.

for a specified thermal velocity V_T. We applied the method described in [16], which uses numbers with inverted orders.

The proton trajectories were calculated using a second-order implicit combined algorithm that precisely conserves the total energy (Hamiltonian) of the particle:

$$H(\mathbf{x},\mathbf{v}) = e\varphi(\mathbf{x}) + m_0c^2\gamma(\mathbf{v}), \qquad (5)$$
$$\gamma(\mathbf{v}) = (1-\mathbf{v}^2/c^2)^{-1/2},$$

where e, m_0, and c are the charge and rest mass of the proton and the speed of light. The algorithm operates in accordance with the following scheme. If a particle is not magnetized, the Newton–Lorentz equations of motion are numerically integrated using the relativistic generalization of the implicit scheme of Borodachev [17], which is accurate to second order.

If a particle is magnetized, the motion of its driving center is calculated by numerically integrating the drift equations of motion using a second-order implicit scheme, while the Larmor rotation about the driving center is described by the relevant analytical formulas. If the electrical-drift velocity $\mathbf{V}_E = c[\mathbf{E}\times\mathbf{B}]/|\mathbf{B}|^2$ is small compared to the velocity of the particle $\mathbf{v}$, $|\mathbf{V}_E|/|\mathbf{v}| \ll 1$, we use the drift equations for a weak electric field in a form that is suitable in both the relativistic and classical cases; if this condition is not satisfied and the particle is classical ($|\mathbf{v}|/c \ll 1$), we use the drift equations for a strong electrical field. At each step in the phase space $R^6_{\mathbf{x},\mathbf{v}}$ along the vector $\nabla_{\mathbf{x},\mathbf{v}}H(\mathbf{x},\mathbf{v})$, we project onto the supersurface

$$H(\mathbf{x},\mathbf{v}) = H(\mathbf{x}^0,\mathbf{v}^0) = \text{const}, \qquad (6)$$

where $(\mathbf{x}^0,\mathbf{v}^0)$ are the initial phase variables of the particle when it is generated, thus leading to the exact conservation of the total energy of the particle. The exact definition of a magnetized particle and a detailed description of the drift algorithm used here are presented in [18].

When determining the energy spectrum, we chose an energy step ΔW and have found the distribution of the energy gained by the particles up to the time they leave the parallelepiped Π. To shorten the computation time, only the trajectories of particles leaving the layer $\Pi(1)$ were calculated. The trajectories of particles leaving the layer $\Pi(k)$ were determined by shifting the trajectories for the particles leaving $\Pi(1)$ upward a distance $(k-1)L_g$ along the x_3 axis. The computations were carried out as follows. We calculated the trajectory for each particle beginning its motion in the layer $\Pi(1)$ as long as it was located in the parallelepiped Π. Each time a particle crossed the boundary of some layer $\Pi(k)$, unity was added to the spectral interval of width ΔW in which the energy W of the particle fell; i.e., this determined the corresponding contribution to the energy spectrum $\Delta N_{j(W)}$:

$$N_{j(W)} \longmapsto N_{j(W)} + \Delta N_{j(W)}. \qquad (7)$$

The kinetic energy $W = m_0c^2[\gamma(\mathbf{v})-1]$ and number of the energy interval $j(W) = W/\Delta W$ were calculated for this purpose. The contribution to the spectrum was calculated as follows. If a particle leaves the side of the layer $\Pi(k)$, i.e., it leaves Π altogether, then $\Delta N_{j(W)} = \nu_{||} + 1 - k$; if the particle leaves the layer $\Pi(k)$ in the forward direction, i.e., it enters the layer $\Pi(k+1)$ when $k = 1,\ldots,\nu_{||}-1$ or leaves Π when $k = \nu_{||}$, then $\Delta N_{j(W)} = 1$. This process yields an accumulated spectral dataset $\{N_j\}$, where N_j is the number of particles with energy in the half-interval $j\Delta W \le W < (j+1)\Delta W$. The periodicity in the coordinate x_3 introduced when the particles are generated enabled us to calculate the trajectories of only those particles leaving the layer $\Pi(1)$, making it possible to carry out the simulations on a personal computer, but also introduced certain nonphysical oscillations in the computed spectrum.

In the simulation, we used the values for the model constants in (1)–(4) $L_g = 80$ m, $L_{||} = 10^8$ m, $L_\perp = 5040$ m (i.e., $\nu_{||} = 1\,250\,000$, $\nu_\perp = 63$), $E_0 \approx 4\times 10^4$ V/m, $B_0 \approx (1\text{–}2)\times 10^2$ G, $V_T = 0.001$ c, and $\Delta W = 0.001\times m_0c^2 \approx 0.938$ MeV. The maximum time for the particles leaving the layer $\Pi(1)$ to be located in the parallelepiped Π was less than 10^{-3} s in all cases, which is three orders of magnitude smaller than the characteristic time for the variation of the fields, which is of the order of 1 s.

Figure 6 presents the energy spectrum for the case $E_0 = 4\times 10^4$ V/m, $B_0 = 10^2$ G on a semilog scale.

Like the spectrum of the fast component, this spectrum displays a well-defined exponential character, indicating the acceleration of ions by an electric field during the flare.

5. DISCUSSION

The analysis of the neutron-monitor signals revealed the presence of both the fast and delayed components of the RSPs. The fast component has an exponential energy spectrum. A close to exponential spectrum was also observed in the early stages of events for which results have been published previously, such as the event of September 29, 1989 [13, 19]. The spectrum for the event of May 7, 1978, which was composed exclusively of the fast component of the RSPs, was likewise exponential [1]. Various authors have pointed out that an exponential energy spectrum is characteristic of acceleration by the electric field that arises during the reconnection of oppositely directed magnetic fields in a plasma.

It is concluded in [13, 20] that the spectrum of the fast component results from acceleration by electric fields in rapidly reconnecting coronal current sheets during flares. The expression obtained in [21] was used in [20] to calculate the energy spectrum of protons accelerated by this mechanism under a number of simplifying assumptions. Such calculations carried out for energies below 100 MeV in [11] led the authors to conclude that the spectrum of protons accelerated by such an electric field is a power-law rather than an exponential function. However, the small difference between exponential and power-law spectra at such low energies casts doubt on the correctness of this conclusion.

We especially wish to emphasize that, in all theoretical studies, including [9, 10], models with a constant magnetic configuration with a central X line have been considered, rather than the magnetic field of a current sheet in which there is rapid reconnection during a flare. Given the contradictory results yielded by theoretical analyses carried out by essentially the same group [9–11], it was clearly necessary to perform independent calculations.

The computational power of modern computers makes them capable of carrying out reasonably accurate computations of the particle spectrum that are consistent with our current understanding of the nature of magnetic reconnection and take into account the available observational data. However, this requires overcoming a number of technical problems. Our computations were also carried out assuming a stationary field, but for energies to 1000 MeV, enabling us to compare our results with spectra derived from neutron-monitor measurements and thereby to obtain results that are qualitatively different from those of [11]. These results provide evidence for the acceleration of particles by an electric field, as occurs in the electrodynamical model for a flare. Further, we plan to carry out computations for a more realistic current sheet in which the electric field is generated during rapid magnetic reconnection. We propose to construct the initial conditions for these computations using the results of numerical MHD simulations of a current sheet above a real active region, based on the real photospheric perturbations observed before a flare [14, 15].

The spectrum of the slow component is a power law extending from 10^2 to 10^4 MeV, possibly testifying to the presence of a single source for the slow component in this energy interval. The slow-component spectrum is probably formed by turbulence, as well as the diffusion of relativistic particles captured by magnetic fields due to scattering on inhomogeneities of the magnetic field. Particle acceleration in shock fronts or in magnetic traps during variations of the fields in postflare loops is also possible [22].

6. CONCLUSIONS

(1) Analysis of solar relativistic protons for several flares in the 23rd solar cycle demonstrates that the spectrum of the particles arriving at the front of the flux of flare cosmic rays (the fast component) displays a well-defined exponential character. The spectrum of the particles that are detected 20–30 min after the arrival of the RSP front can be described well by a power law.

(2) Numerical simulation shows that the spectrum of particles accelerated by an electric field applied along a neutral line of the magnetic field is exponential.

(3) Our results support the flare electrodynamical model, in which protons are accelerated during the flare by the electric field generated along the neutral X line of the magnetic field during reconnection. The formation of the delayed component of the solar cosmic rays produced is associated with the generation of turbulence or the gradual escape of accelerated particles that are captured by magnetic fields. The diffusion of particles due to scattering on magnetic inhomogeneities in the interplanetary magnetic field can also affect the shape of the delayed component's spectrum.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16987) and the Federal Science and Technology Program in Astronomy.

REFERENCES

1. M. A. Shea and D. F. Smart, Space Sci. Rev. **32**, 251 (1982).
2. E. V. Vashenyuk, Yu. V. Balabin, and B. B. Gvozdevsky, in *Proc. 28th Int. Cosmic Ray Conf., Tsukuba, Japan, 2003*, Vol. 6, p. 3401.
3. A. I. Podgorny and I. M. Podgorny, Sol. Phys. **139**, 125 (1992).
4. A. I. Podgorny and I. M. Podgorny, Izv. Ross. Akad. Nauk, Ser. Fiz. **61**, 1067 (1997).
5. A. I. Podgorny and I. M. Podgorny, Astron. Zh. **78**, 71 (2001) [Astron. Rep. **45**, 60 (2001)].
6. R. P. Lin, S. Krucker, G. J. Hurford, *et al.*, Astrophys. J. Lett. **595**, L69 (2003).
7. S. Krucker, G. J. Hurford, and R. P. Lin, Astrophys. J. Lett. **595**, L103 (2003).
8. E. Hiei and A. J. Hundhausen, in *IAU Colloq. No. 153: Magnetodynamic Phenomena in the Solar Atmosphere, Prototypes of Stellar Magnetic Activity,* Ed. by Y. Uchida, T. Kosugi, and H. S. Hudson (Kluwer, Dordrecht, 1996), p. 125.
9. S. V. Bulanov and P. V. Sasorov, Astron. Zh. **52**, 763 (1975) [Sov. Astron. **19**, 464 (1975)].
10. S. V. Bulanov and F. Chap, Astron. Zh. **65**, 837 (1988) [Sov. Astron. **32**, 436 (1988)].
11. K. Mori, J. Sakai, and J. Zhao, Astrophys. J. **494**, 430 (1998).
12. E. V. Vashenyuk, O. V. Mingalev, and B. B. Gvozdevskii, Izv. Ross. Akad. Nauk, Ser. Fiz. **67**, 455 (2003).
13. E. V. Vashenyuk, L. I. Miroshnichenko, and B. B. Gvozdevsky, Nuovo Cimento **23**, 285 (2000).
14. I. A. Bilenko, A. I. Podgorny, and I. M. Podgorny, Sol. Phys. **207**, 323 (2002).
15. A. I. Podgorny, I. M. Podgorny, and I. A. Bilibenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **67**, 406 (2003).
16. Ch. K. Birdsall and A. B. Langdon, *Plasma Physics via Computer Simulations* (McGraw-Hill, New York, 1985; Énergoatomizdat, Moscow, 1989).
17. L. V. Borodachev, Zh. Vychisl. Mat. Mat. Fiz. **31**, 934 (1991).
18. L. V. Borodachev, I. V. Mingalev, and O. V. Mingalev, Zh. Vychisl. Mat. Mat. Fiz. **43**, 467 (2003).
19. E. V. Vashenyuk, V. V. Pchelkin, and L. I. Miroshnichenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **65**, 350 (2002).
20. J. Peres-Peraza, K. A. Gal'egos, E. V. Vashenyuk, and L. I. Miroshnichenko, Geomagn. Aeron. **32**, 2 (1992).
21. J. Perez-Peraza, M. Galvez, and A. Lara, in *Proc. 15th Int. Cosmic Ray Conf., Plovdiv, Bulgaria, 1977*, Vol. 5, p. 23.
22. B. V. Somov and S. A. Bogachev, Pis'ma Astron. Zh. **29**, 701 (2003) [Astron. Lett. (2003)].

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 10, 2005, pp. 847–857. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 10, 2005, pp. 950–960.

A Gravitational–Tidal Mechanism for the Earth's Polar Oscillations

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

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

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

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

Received December 12, 2004; in final form, February 17, 2005

Abstract—Perturbed, rotational–oscillational motions of the Earth induced by the gravitational torques exerted by the Sun and Moon are studied using a linear mechanical model for a viscoelastic rigid body. A tidal mechanism is identified for the excitation of polar oscillations, i.e., for oscillations of the angular-velocity vector specified in a fixed coordinate frame, attributed to the rotational–progressive motion of the barycenter of the Earth–Moon "binary planet" about the Sun. The main features of the oscillations remain stable and do not change considerably over time intervals significantly exceeding the precessional period of the Earth's axis. A simple mathematical model containing two frequencies, namely, the Chandler and annual frequencies, is constructed using the methods of celestial mechanics. This model is adequate to the astrometric measurements performed by the International Earth Rotation Service (IERS). The parameters of the model are identified via least-squares fitting and a spectral analysis of the IERS data. Statistically valid interpolations of the data for time intervals covering from several months to 15–20 yr are obtained. High-accuracy forecasting of the polar motions for 0.5–1 yr and reasonably trustworthy forecasting for 1–3 yr demonstrated by observations over the last few years are presented for the first time. The results obtained are of theoretical interest for dynamical astronomy, geodynamics, and celestial mechanics, and are also important for astrometrical, navigational, and geophysical applications. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Let us first present some brief historical and preliminary methodical notes. Numerous astrometric studies have been based on the dynamical theory of the Earth's rotation about its center of mass [1–7]. It is well known from observations (regular measurements have been performed since the end of the 19th century) that the Earth's rotational axis changes its orientation in time, with respect to both fixed and inertial coordinate frames. The rather complicated polar oscillations contain components with very different frequencies and amplitudes. For example, small variations of the angular-velocity vector in a reference frame fixed to the Earth contain a main component with an amplitude reaching $0.20''$–$0.25''$ and a period of approximately 430–440 sidereal days, which was discovered by Chandler in 1891. The significant difference between the Chandler period and the Euler precessional period of 305 days, specified by the classical theory of rigid-body rotation for a nondeformable Earth, required explanation, which was sought and partially found in studies by numerous researchers, such as Newcomb, Poincaré, Jeffreys, Love, Melchior, Munk, Macdonald, Sludskii, and Molodenskii [2–6], based on models for the deformable Earth. This motion is called the free nutation of the deformable Earth, or the Chandler wobble.

It was found (also by Chandler) that there was an appreciable component with an amplitude of $0.07''$–$0.08''$ and a period of one year (about 365.25 sidereal days). The observed polar oscillations have the character of beating. The trajectory of the pole on the Earth's surface takes the form of a spiral that rolls up and unrolls with a period of about six years (see below).

The trajectory of the pole and forecasting of this trajectory are of considerable interest both in their own right and for various applications. Constructing a high-accuracy theoretical model for the rotation of the deformable Earth, identifying parameters of the model using the data of the International Earth Rotation Service (IERS), and trustworthy forecasting of the pole motion are very important for navigation over long time intervals [8] and for a number of astrometrical, geodynamical, and geophysical applications [1–7].

A simple mechanical model for a viscoelastic rigid body [9] can be used to describe the rotation of the

1063-7729/05/4910-0847$26.00

deformable Earth and the oscillations of the pole (the North Pole for definiteness). To take account of gravitational tidal effects, we will assume the Earth to be axially symmetric and to have two layers: a completely rigid spherical core and a viscoelastic mantle. Of course, one can consider more complex models. However, using more complex models for the Earth's figure is not justified, since the relevant geometrical and physical parameters cannot be determined with the required accuracy and completeness via statistical studies of the available indirect measurements. We will adhere to the idea that the complexity of the model should strictly correspond to the formulated problem and to the accuracy of the observational data. When constructing a model for the polar oscillations, we can restrict our treatment to a limited number of mean (integrated) characteristics of the inertia tensor. Comparison with measurements and further analysis demonstrate that these adopted simplifications are justified.

For a preliminary study of the polar motions and their evolution due to perturbing moments, we consider the spatial case of the two-body problem [9, 10]. We assume that the center of mass of the deformable planet (the Earth) and its point satellite (the Moon) undergo some mutual translational–rotational motion about their mutual center of mass (barycenter), which in turn moves along an elliptical orbit around the Sun.

An asymptotic analysis of the equations of motion in osculating action–angular variables enables us to determine (in a quasi-static approximation) some stable characteristics of the rotational–oscillatory motion of the deformable Earth about the center of mass. We first determine refined periods (frequencies) for the axial rotation and Chandler oscillation and compare these with the results of a spectral analysis [1, 2]. We then find estimates for the amplitudes of the free oscillations of the angular-velocity vector in a fixed coordinate frame and compare these with observations.

Using the kinematic Euler and dynamical Euler–Liouville equations, we construct a first-approximation model for the Chandler and annual polar oscillations induced by the gravitational tidal forces applied by the Sun and Moon. Finally, we present numerical estimates of the parameters of the motion derived via a least-squares fit to daily measurement data and compare the trajectory and forecasted motion of the Earth's pole with the astrometric IERS data [1] (see also http://hpiers.obspm.fr/eoppc/eop/eoppc04/eoppc04-xy.gif).

2. PRELIMINARY ESTIMATES FOR FREE NUTATION OF THE EARTH'S POLE

Let us determine the oscillations of the angular-velocity vector of the deformable Earth in some fixed ("frozen") reference coordinate frame [1, 5]. To study free nutation (the Chandler component of the polar motion), it is convenient to turn from the Andoyer variables to the action–angular variables I_j and w_j $(j = 1, 2, 3)$ [9]. The action–angular variables specify the Hamiltonian variables. The following dynamical parameters κ and λ describe the Earth's rotational motions:

$$\kappa^2 = \frac{C^*(A^* - B^*)}{A^*(B^* - C^*)}, \quad \lambda^2 = \kappa^2 \frac{2EC^* - I_2^2}{I_2^2 - 2EA^*}. \quad (1)$$

Here, A^*, B^*, and C^* are the effective main central moments of inertia taking into account deformations of the "frozen" Earth due to a compound motion that includes the Earth's rotation about its axis and its motion about the Earth–Moon barycenter. The constant E in (1) is the integral of the kinetic energy for an intermediate (rotational) motion of the deformable Earth, while $I_2 = |G|$ is the absolute value of the Earth's kinetic moment. The relations between the action–angular variables I_j, w_j $(j = 1, 2, 3)$ and the canonical Andoyer variables L, G, $G_{\xi 3}$, φ_j are expressed through elliptical Jacobi functions and their integrals [9].

We can express the Routh functional for the intermediate model through the action–angular variables and construct trajectories in the phase space of I_j, w_j $(j = 1, 2, 3)$. In particular, the Routh functional R_0 averaged over the fast angular variables φ_2 and ϑ describing the rotation and orbital motions takes the form [9]

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

The general preliminary solution of the averaged problem corresponding to R_0 (2) takes the form

$$I_i(t) = I_i^0 = \text{const}, \quad w_3(t) = w_3^0, \quad (3)$$
$$w_{1,2}(t) = n_{1,2}t + w_{1,2}^0, \quad w_i^0 = \text{const};$$
$$n_1 = -\frac{\pi}{2}\frac{I_2}{A^*}\frac{\mu^*\kappa}{\kappa_* \chi K(\lambda)}, \quad \kappa_* = \sqrt{1 + \kappa^2},$$
$$K = \frac{\pi}{2} + \frac{\lambda^2}{4} + O(\lambda^4),$$
$$n_2 = \frac{I_2}{A^*}\left(\frac{A^*}{C^*} + \mu^* \frac{\Pi(\pi/2, \kappa^2, \lambda)}{K(\lambda)}\right),$$
$$\chi = \kappa\sqrt{1 + \varepsilon^2}, \quad \Pi = \frac{\pi}{2\kappa_*} + O(\lambda^2).$$

Here, I_j^0 and w_j^0 are the initial conditions determined through the Andoyer variables for $t = 0$. The phases w_1, w_2 and frequencies n_1, n_2 correspond to the polar Chandler motion and diurnal rotation of the deformable Earth, respectively.

Thus, in its "intermediate motion," the deformable Earth rotates uniformly in the fields of the centrifugal-inertial and lunar-gravitational forces. In the case of a completely rigid model planet, there is a regular Euler–Poinsot precession. For our case of the deformable Earth, expressions (3) also describe a regular precession, although the angular velocities of the precession and rotation exhibit some comparatively small variations due to perturbing effects.

For motion perturbed by the dissipative effects of the Earth's viscoelastic mantle, there is a regular precession with slowly varying parameters, i.e., an evolution of the slow variables, which must be studied using asymptotic methods of nonlinear mechanics. The solution (3) is the generating equation for the averaging method.

The estimate of [1, 5] for λ for the Earth–Moon system indicates that $\lambda \ll 1$ $(\lambda^2 \sim 10^{-14})$, i.e., μ^* is small. This enables us to appreciably simplify all the formulas (3) and express them in terms of algebraic and trigonometric functions. We find in a first approximation

$$\frac{\lambda^2}{\kappa^2} = \varepsilon^2 = \frac{2\mu^*}{\kappa_*}. \tag{4}$$

Substituting λ^2 (4) into expressions (3) for the frequencies n_1 and n_2 of the intermediate motion yields

$$\dot{w}_1 = n_1 = -\frac{I_2}{A^*}\frac{2\mu^*}{\kappa_*}\left(1 - \frac{2+\kappa_*^2}{2\kappa_*}\mu^*\right), \tag{5}$$

$$\dot{w}_2 = n_2 = -\frac{I_2}{A^*} \times \left(\frac{A^*}{C^*} + \frac{\mu^*}{\kappa_*}\left[1 + \frac{\mu^*}{2\kappa_*}(2\kappa^* - 2 - \kappa^2)\right]\right).$$

We can calculate the theoretical angular velocity and period of the oscillations $T_1^* = 2\pi/n_1 \approx 430$ days using formula (5) for n_1. This period agrees well with the experimental measurements, which give $T_1 \approx 420$–440 sidereal days, known as the Chandler period for the Earth's polar oscillations [1–7].

We define the polar motion as the angular displacement of the rotational axis determined in a coordinate frame fixed to the planet's body. The components of the angular-velocity vector are expressed via the phase w_1 in the form of expansions in the small parameter $\varepsilon \sim 10^{-6}$. With the relative error of $O(\varepsilon^2) \sim 10^{-12}$, the angular-velocity components are [9]

$$\omega_1 = \frac{\lambda}{\chi}\frac{I_2}{A^*}\mathrm{cn}(u,\lambda) = \varepsilon\frac{I_2}{A^*}\cos w_1 + O(\varepsilon^3), \tag{6}$$

$$\omega_2 = \frac{\lambda}{\chi}\frac{I_2}{B^*}\mathrm{sn}(u,\lambda) = \varepsilon\sqrt{1+\kappa^2}\frac{I_2}{B^*}\sin w_1 + O(\varepsilon^3),$$

$$\omega_3 = \frac{\kappa}{\chi}\frac{I_2}{C^*}\mathrm{dn}(u,\lambda) = \frac{I_2}{C^*} + O(\varepsilon^2),$$

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

Taking into account the main terms of the expansions (6), the angular coordinates (x_c, y_c) for the free nutation (the Chandler component of the motion), the angle α between the axis of the figure and the rotational axis, and the linear coordinates X_c, Y_c in the plane tangent to the geoid take the form

$$x_c = \frac{\omega_1}{\omega} \approx \frac{C^*}{A^*}\frac{\lambda}{\kappa}\cos w_1, \tag{7}$$

$$y_c = -\frac{\omega_2}{\omega} \approx \frac{C^*}{B^*}\frac{\lambda}{\kappa}\kappa_*\sin w_1,$$

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

$$\cos\alpha = \frac{\omega_3}{\omega} \approx 1 - \frac{\varepsilon}{2} \times \left(\frac{C^{*2}}{A^{*2}}\cos^2 w_1 + (1+\kappa^2)\frac{C^{*2}}{B^{*2}}\sin^2 w_1\right),$$

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

In a first approximation in ε, the polhode corresponding to the free nutation of the Chandler period is an ellipse with a fairly small eccentricity, $e \approx 0.005$. The IERS data [1] support the theoretical estimates (7). An analysis of the effects of the tidal moments of dissipative forces shows that both the period and amplitude of the oscillations vary insignificantly over periods greatly exceeding the precessional period of the Earth's axis.

3. FIRST-APPROXIMATION MODEL FOR THE POLAR OSCILLATIONS

To construct a model for the rotational motion about the center of mass, let us express the equations in the form of the classical dynamical Euler–Liouville equations for the varying inertia tensors J [1–6, 10, 11],

$$J\dot{\omega} + \omega\times J\omega = M, \quad \omega = (p,q,r)^T, \tag{8}$$

$$J = J^* + \delta J, \quad J^* = \mathrm{const},$$

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

$$||\delta J|| \ll ||J^*||.$$

Here, ω is the angular-velocity vector in the coordinate frame [5] which is fixed to the Earth (for epoch 1900) and approximately coincides with the main central inertia axes for J^* of the "frozen" Earth taking into account the equatorial bulge [1–6, 9]. Additional perturbing terms arising during the differentiation of the kinetic moment of the deformable Earth [9, 10] are added to the rather complex vector of the perturbing moment M. We expect that small variations of the inertia tensor δJ can contain various harmonics associated with the solar and lunar diurnal gravitational tides, and possibly some other harmonics (annual, semiannual, monthly, semidiurnal, and so on [1–5]). Gravitational tidal effects with the annual and near-Chandler periods are thought to be the main external sources for the perturbing moments M inducing the nutational oscillations (see below).

Adding a possible term of the type $\dot{J}\omega$, or rather of $\dot{G} - J\dot{\omega} - \omega \times J\omega$, does not improve the first-approximation model. Attempts to estimate these terms accurately taking into account some geophysical effects are difficult and do not give satisfactory results. Analyses of these effects in the literature have often not been constructive, since they do not determine the moments of the forces (their amplitudes, directions, and frequency and phase characteristics).

Let us consider first a simple mechanism for the excitation and maintenance of nutational oscillations with the annual period. The kinematic Euler equations determining the orientation of the fixed axes in the orbital coordinate frame take the form [8, 9]

$$\dot{\theta} = p\cos\varphi - q\sin\varphi - \omega_0(\nu)\sin\psi, \qquad (9)$$
$$\dot{\nu} = \omega_0(\nu) = \omega_*(1 + e\cos\nu)^2,$$
$$\dot{\psi} = \frac{p\sin\varphi + q\cos\varphi}{\sin\theta} - \omega_0(\nu)\cot\theta\cos\psi,$$
$$e = 0.0167,$$
$$\dot{\varphi} = r - (p\sin\varphi + q\cos\varphi)\cot\theta + \omega_0(\nu)\frac{\cos\psi}{\sin\theta}.$$

Here, $\nu(t)$ is the true anomaly, e the orbital eccentricity, and ω_* a constant determined by the gravitational and focal parameters. For the polar motions, the terms proportional to ω_0 in (9) are much greater than the terms with p and q (by approximately a factor of 300) and become dominant for $\dot{\theta}$ and $\dot{\varphi}$. This important feature was not considered earlier, and these terms were inappropriately neglected (the orbital and rotational motions were separated) [1–6].

4. ANNUAL COMPONENT OF THE EARTH'S POLAR OSCILLATIONS

The expressions for the components of the solar gravitational moment take the form [11]

$$M_q^s = 3\omega^2[(A^* + \delta A - (C^* + \delta C))\gamma_r\gamma_p \qquad (10)$$
$$+ \delta J_{pq}\gamma_r\gamma_q + \delta J_{pr}(\gamma_r^2 - \gamma_p^2) - \delta J_{rq}\gamma_p\gamma_q],$$
$$\omega = \omega_*(1 + e\cos\nu)^{3/2},$$
$$\gamma_p = \sin\theta\sin\varphi, \quad \gamma_q = \sin\theta\cos\varphi, \quad \gamma_r = \cos\theta.$$

To calculate $M_{p,q}^s$, we must cyclically rearrange the indices p, q, and r in (10). Expressions (10) show that the annual polar oscillation can be attributed to a term containing the products of the directional cosines $\gamma_p\gamma_r$ and $\gamma_p\gamma_r$. To calculate these in a first approximation (for $p, q = 0$), we integrate Eqs. (9):

$$r = r^0, \quad \varphi \approx rt + \varphi^0, \quad \nu \approx \omega_* t + \nu^0, \qquad (11)$$
$$\cos\theta(\nu) = a(\theta^0, \psi^0)\cos\nu, \quad \theta(0) = \theta^0 = 66^\circ 33',$$
$$0.4 \le a \le 1, \quad 0 \le \psi^0 \le 2\pi,$$
$$\cos\theta\sin\theta = b(\theta^0, \psi^0)\cos\nu + d\cos 3\nu + \ldots,$$
$$0.4 \le b \le \frac{4}{3\pi}, |d| \ll 1.$$

The second and higher harmonics on ν are neglected, since their contributions are a factor of 10^2–10^3 smaller than that of the main harmonic. The quantity $B^* - A^*$ is also considerably smaller than $C^* - A^*$ (by approximately a factor of 160). Estimating the terms for p and q in (8), using (11), and averaging over the fast phase variable φ, we obtain the simplified analytical model

$$\dot{p} + N_p q = \kappa_q r^2 + 3b\omega_*^2\chi_p^s\cos\nu, \qquad (12)$$
$$N_{p,q} \approx N = \frac{2\pi}{T_1} \approx (0.84\text{–}0.85)\omega_*,$$
$$\dot{q} - N_q p = -\kappa_p r^2 - 3b\omega_*^2\chi_q^s\cos\nu,$$
$$p(0) = p^0, \quad q(0) = q^0.$$

Here, κ_p and κ_q are the averages of $\delta J_{pr}/B^*$ and $\delta J_{qr}/A^*$, which can be slowly varying. The quantities χ_p^s and χ_q^s result from averaging the coefficients of $\cos\nu$ in the components of the solar gravitational moment over the variable φ. Due to the axial symmetry of the problem, we can assume that $\chi_{p,q}^s = \chi$, where χ is determined from observations. These quantities are associated with the diurnal tides.

The lunar gravitational moments with the monthly period of 27.55 days are neglected because of their relatively weak effect on the annual nutational oscillations (due to their considerably different frequencies). These effects become appreciable for more detailed investigations of the extrema of the polar deviations

in each component. The effects of the monthly perturbations are especially clear during beating (the minimum amplitudes of the oscillations, see below).

The right-hand terms of (12) contain an explicit harmonic of the annual period, which can explain the excitation of the nutational oscillations detected in the IERS data. The sensitivity of the coefficients $\kappa_{p,q}$ is higher than the sensitivity of $\chi^s_{p,q}$ by five orders of magnitude. This has led to attempts at geophysical–seasonal treatments of the annual component. However, any explicit regular mechanism for the annual perturbation with the required amplitudes of $M_h \sim 10^{20}$ kg m^2 s^{-2} due to internal geophysical factors (atmospheric effects, oceanic flows, seasonal phenomena on the Earth's surface, etc.) proves to be inconsistent from a mechanical point of view. Analyses of the frequency distribution of the annual oscillations (the sharpness of the peak of the spectral density of the time series) and of the phase shifts for various processes show that geophysical interpretations are not sound [2].

The ocean and atmosphere are also exposed to considerable gravitational tidal effects that lead to annual oscillations of the Earth's pole. These contributions can be estimated in the hydrodynamic models. These contributions most likely explain the synchronism in the variations of the dynamical characteristics of the atmosphere, ocean, and mantle and of the polar oscillations.

5. THE CHANDLER COMPONENT OF THE POLAR OSCILLATIONS

As the central body of the solar system, the Sun exerts the dominant effect on the orbital motion of the planets. The perturbing influences of the planets on each other and their satellites is rather weak. This conclusion follows from analysis and averaging of the osculating orbital elements. The planetary motions are quite close to Keplerian.

The motion of the Earth–Moon system differs significantly from the typical planetary motion. The dynamical characteristics of the system are very specific in comparison with the other planets [2–5], and a number of dynamical problems treat this system as a binary planet [9, 10]. Studies of the translational–rotational motion of the system have used various formulations, with the following being most typical.

(1) Classical gravitational theory describes the orbital motion of the Earth–Moon system using point-mass models.

(2) The translational–rotational motion of the Earth and Moon as rigid bodies is studied using the classical methods of dynamical theory. The equations of motion in either Euler or Andoyer variables are solved using perturbation methods considering small parameters. However, this theory cannot explain a number of phenomena, in particular, the polar oscillations.

(3) Modern high-accuracy studies of rotations and oscillations about the center of mass take into account elastic deformations of the complex figures of the Earth and Moon occurring due to the gravitational tides.

(4) Evolutionary theories examine elastic-dissipative effects taking into account both gravitational tides and irregular geophysical perturbations. These studies are presently still under intense development.

When dissipation is not included, the gravitational tidal forces have a potential that is a sum of harmonics. In our case, when the Moon can be treated as a gravitating point or sphere, the potential U_M takes the form [2, 3]

$$U_\mathrm{M} = -\frac{1}{4}Kg\frac{r^2}{R_\mathrm{E}}(1-3\cos^2\theta)\sum_i^{(0)} C_i\cos A_i \quad (13)$$
$$-\frac{1}{2}Kg\frac{r^2}{R_\mathrm{E}}\sin 2\theta\sum_i^{(1)} C_i\cos(A_i+\varphi)$$
$$-\frac{1}{2}Kg\frac{r^2}{R_\mathrm{E}}\sin^2\theta\sum_i^{(2)} C_i\cos(A_i+2\varphi),$$
$$K = \frac{3}{2}\frac{m_\mathrm{M}}{m_\mathrm{E}}\left(\frac{R_\mathrm{E}}{R_\mathrm{EM}}\right)^3 = 0.843\times 10^{-7},$$
$$0 \le r < R_\mathrm{E}.$$

Here, θ and φ are the geographical coordinates of the point, r is its distance from the center of mass, R_E is the mean radius of the Earth ($R_\mathrm{E} \simeq 6.38\times 10^6$ m), R_EM is the mean distance from the Earth's center of mass to the Moon, and g is the gravitational acceleration. The angular variables A_i are integer linear combinations of the six parameters τ_0, $l_\mathrm{M,S}$, $p_\mathrm{M,S}$, and Ω_M. Here, $\tau_0 = t - l_\mathrm{M} + l_\mathrm{S}$ is the mean lunar Greenwich time and t is the mean solar Greenwich time. The parameters l_M and l_S are the mean lunar and solar longitudes, which vary with periods of 27.55 and 365.25 sidereal days, respectively. p_M is the mean longitude of the lunar perigee, which varies with a period of 8.85 yr, and p_S is the mean longitude of the solar perigee, which varies with a period of 25.700 yr. The parameter Ω_M determines the longitude of the ascending lunar node and varies with a period of 18.61 yr.

The sums $\sum_i^{(j)}$ take account of long-period, diurnal, and semidiurnal gravitational tidal lunar effects. The potential components corresponding to the diurnal tides are of the most interest for analysis of the

polar oscillations. We take into account the components with six-year modulations due to the lunar–solar precession. For the components along the nonrotating axes of the Earth in the rotating barycentric coordinate frame, we obtain the expressions

$$\tilde{M}_{p,q} = \kappa^{\mathrm{M}}_{p,q}\cos\psi_{\mathrm{M}}, \tag{14}$$
$$\psi_{\mathrm{M}} = \Omega_{\mathrm{M}} + p_{\mathrm{M}} = \omega_{\mathrm{M}}\tau + \psi^0_{\mathrm{M}},$$
$$\Omega_{\mathrm{M}} = \frac{2\pi}{18.61}\tau + \Omega^0_{\mathrm{M}}, \quad p_{\mathrm{M}} = \frac{2\pi}{8.85}\tau + p^0_{\mathrm{M}},$$
$$\omega_{\mathrm{M}} \simeq \frac{2\pi}{6.00}, \quad \psi^0_{\mathrm{M}} = \Omega^0_{\mathrm{M}} + p^0_{\mathrm{M}},$$

disregarding the orbital motion of the barycenter. Here, $\kappa^{\mathrm{M}}_{p,q}$ are the tidal coefficients, which depend on the fast angular variable $(h_s - l_{\mathrm{M}})$, and τ is the time in years. Taking into account the motion of the orbital barycenter about the Sun in the rotating frame fixed to the Earth, we obtain for the components of the gravitational tidal moment

$$M^{\mathrm{M}}_p = [\tilde{M}_p\cos(h_s - l_{\mathrm{M}}) \tag{15}$$
$$- \tilde{M}_q\sin(h_s - l_{\mathrm{M}})]\cos\theta\sin\theta,$$

$$M^{\mathrm{M}}_q = [\tilde{M}_p\sin(h_s - l_{\mathrm{M}})$$
$$+ \tilde{M}_q\cos(h_s - l_{\mathrm{M}})]\cos\theta\sin\theta.$$

Here, θ is the nutation angle determining the inclination of the axis C^* to the orbital plane, in accordance with (11) [see also (13)].

Substituting (14) and (11) into (15), we obtain multifrequency functions with considerably different periods, which are close to a day, a year, and six years. Applying asymptotic methods for separating the motions (averaging) to approximately calculate the comparatively slow polar oscillations due to the lunar–solar precession yields simplified equations of the form (12)

$$\dot{p} + Nq + \sigma p = f^-_p\cos\alpha^- + f^+_p\cos\alpha^+, \tag{16}$$
$$\dot{q} - Np + \sigma q = f^-_q\cos\alpha^- + f^+_q\cos\alpha^+,$$
$$\alpha^{\mp} = l_s \mp \psi_{\mathrm{M}}.$$

Here, the time is measured in years, and the quasi-constant coefficient σ describes the dissipation. The coefficients $f^+_{p,q} \approx \mathrm{const}$ result from the averaging of the expressions $\kappa^{\mathrm{M}}_{p,q}\cos(h_s - l_{\mathrm{M}})$ and $\kappa^{\mathrm{M}}_{p,q}\sin(h_s - l_{\mathrm{M}})$ over the fast phase $h_s - l_{\mathrm{M}}$. These coefficients are similar to the factors $\chi^s_{p,q}$ in (12). It follows from physical reasoning and observations and measurements that the coefficients in (16) can display small and comparatively slow variations. Fast variations are averaged and virtually do not affect the motion. Assuming the coefficients to be constant over intervals of several years ($\Delta\tau \le 10$ yr), we can construct the required quasi-periodic solution containing the phases $\alpha^{\mp}$.

According to the spectral analysis and calculations, the Chandler period (for the free nutation) is estimated to be in the range $T_1 = 420$–440 days, and the coefficient N (the frequency of the free nutation) varies in the range $N = 0.89$–0.83. The mean period is usually taken to be $T_1 = 433 \pm 2$ days. The frequency of "natural oscillations" for the system (16) is N and does not depend on the dissipation coefficient σ, which is fairly small. The dissipation σ is necessary, since the frequency γ^- (i.e., the rate of change in the phase α^-) is $\gamma^- \approx 5/6 \simeq 0.833$ and can be arbitrarily close to the natural frequency N, according to the estimate (14). In steady oscillations, the amplitude for this frequency can exceed the amplitude for the frequency $\gamma^+ \approx 7/6 \simeq 1.17$ by two to three orders of magnitude. It is quite natural to relate this feature to the difference between the frequencies $\gamma^+ - N$, which considerably exceeds σ. The magnitude of the difference $|N - \gamma^-|$ can be comparable to or smaller than σ. The frequency analysis of the time series also indicates a small peak in the spectral density at the frequency $\gamma^+ \simeq 1.2$. We will neglect this component in our analysis of oscillations of the system (16).

The quasi-stationary Chandler polar oscillations for the frequency $\gamma^- = \gamma$ and amplitudes $f^-_{p,q} = f_{p,q}$ of the external gravitational tidal perturbation are described by the expressions

$$p_{ch} = D^{-1}(D^c_p\cos\alpha + D^s_p\sin\alpha), \tag{17}$$
$$q_{ch} = D^{-1}(D^c_q\cos\alpha + D^s_q\sin\alpha).$$

The coefficients D and $D^{c,s}_{p,q}$ in (17) have the standard form:

$$D^c_{p,q} = \pm 2\sigma\gamma^2 f_{p,q} \tag{18}$$
$$+ (N^2 - \gamma^2 + \sigma^2)(\sigma f_{p,q} \mp \delta f_{q,p})$$
$$\approx 2N^2(\sigma f_{p,q} \pm \delta f_{q,p}),$$

$$D^s_{p,q} = 2\sigma\gamma(\sigma f_{p,q} \mp \delta f_{q,p})$$
$$- \gamma(N^2 - \gamma^2 + \sigma^2)f_{p,q} \approx 2N^2(\delta f_{p,q} \mp \sigma f_{q,p}),$$

$$D = (N^2 - \gamma^2 + \sigma^2) + 4\sigma^2\gamma^2$$
$$\approx 4N^2(\delta^2 + \sigma^2), \quad \delta = \gamma - N, \quad |\delta| \sim \sigma \ll N.$$

The oscillation amplitude and its estimate obtained using (18) are equal to

$$A_{ch} = \sqrt{p^2_{ch} + q^2_{ch}} \approx \frac{1}{2}(f^2_p + f^2_q)^{1/2}(\delta^2 + \sigma^2)^{-1/2}. \tag{19}$$

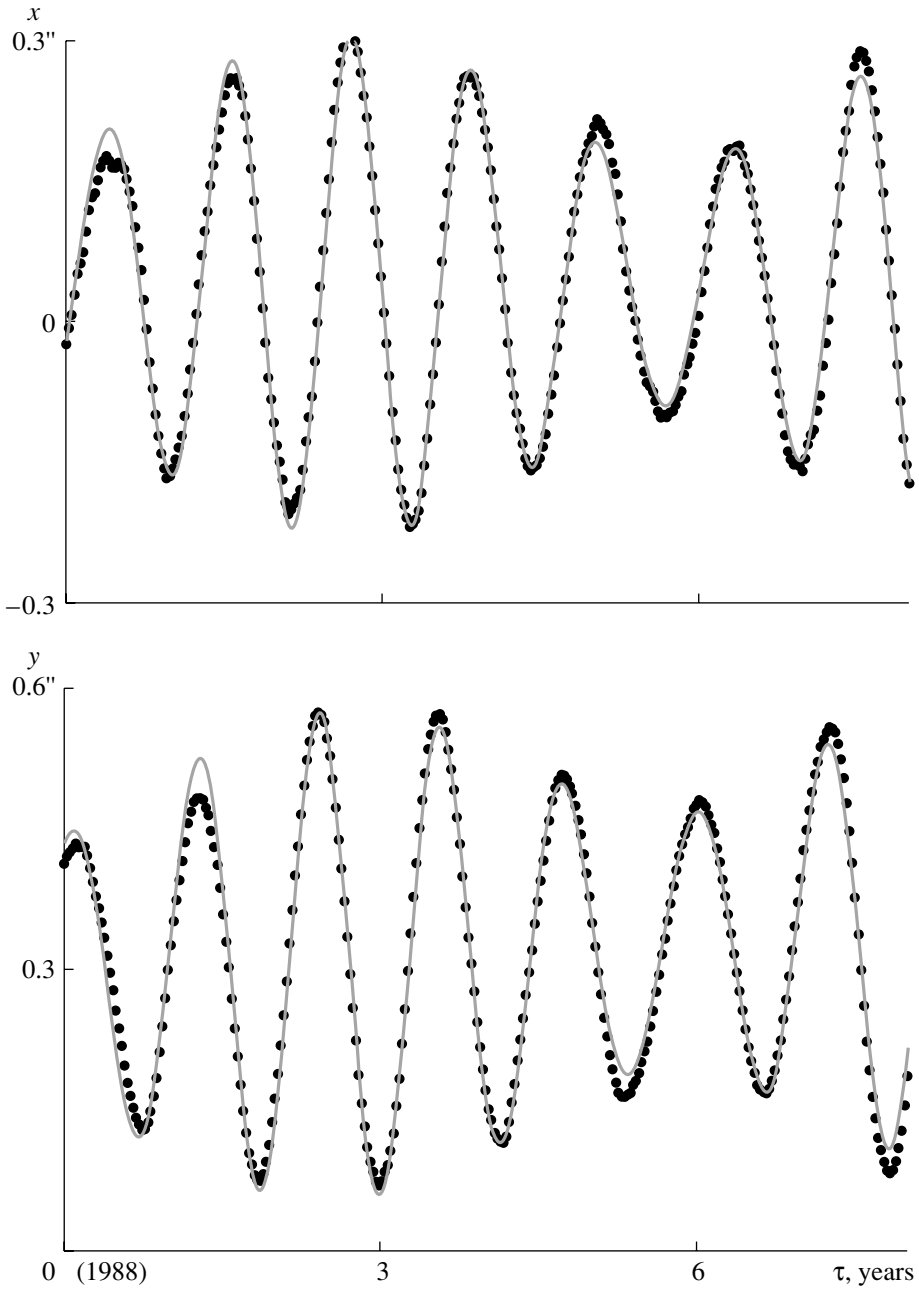

Fig. 1. Interpolation of polar oscillations for the eight-year interval (from 1988 to 1995, inclusive) using the six-parameter model (the points show the IERS data).

Equations (17)–(19) show that the amplitudes of the components and the total amplitude for the Chandler period are very sensitive to the frequency difference δ and dissipation coefficient σ. The presented estimates for small $|\delta|$ and σ indicate a structural property of the oscillations that is analogous to that for the case of free oscillations of the Chandler frequency and annual nutational oscillations [9–11]:

$$D_p^c = D_q^s, \quad D_p^s = D_q^c. \tag{20}$$

Analysis of the IERS data demonstrates that Eqs. (20) are very stable and clearly confirmed by computations [10, 11], although the amplitude of the oscillations depends on the data used. The amplitudes vary in the range $0.25''$–$0.23''$ [1–6], so that we can take the average amplitude to be $0.22''$–$0.23''$, in which case the measurement fluctuations are about $0.02''$, or about 10%.

6. NUMERICAL MODELING

The coefficients $\kappa_{p,q}$ and $\chi_{p,q}$ remain unknown, and we must determine them from the IERS data [1].

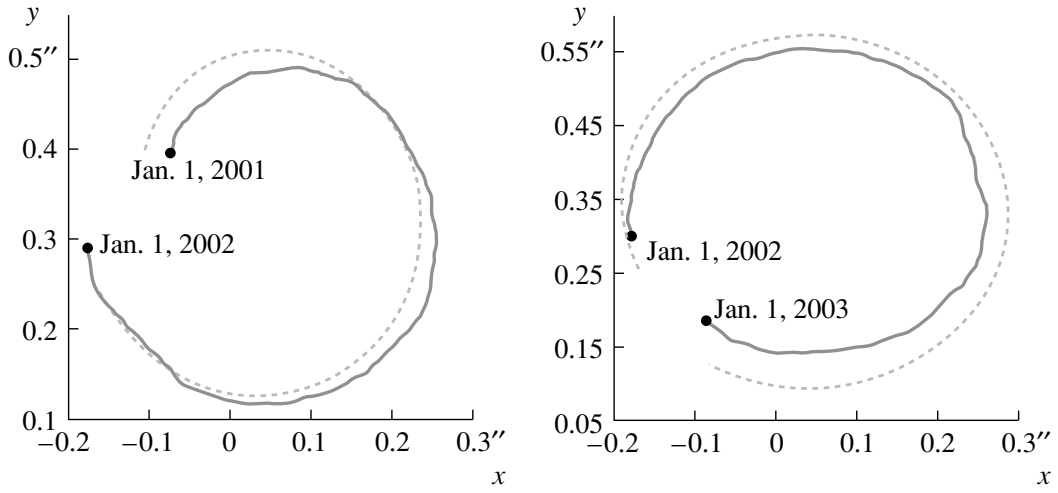

Fig. 2. Annual trajectories of the polar motion over a two-year interval (2001–2002) near the gravitational anomaly of the "parade of planets" in 1999–2000 (right-hand coordinate frame). The solid curves show the IERS data, and the dashed curves show the forecast of the six-parameter model.

Introducing the variables $x(\tau) = p(t)$ and $y(\tau) = q(t)$, where $\tau = t/T_h$ is the time in years, we obtain a solution of (8) in the form

$$x(\tau) = c_x^0 + c_x^1\tau - a_x^c\cos 2\pi N\tau + a_x^s\sin 2\pi N\tau - Nd_x^c\cos 2\pi\tau - d_x^s\sin 2\pi\tau, \tag{21}$$

$$y(\tau) = c_y^0 + c_y^1\tau + a_y^c\cos 2\pi N\tau + a_y^s\sin 2\pi N\tau - Nd_y^c\cos 2\pi\tau + d_y^s\sin 2\pi\tau,$$
$$N = 0.845–0.850.$$

The quantity N is chosen based on a dispersion analysis. We determine the unknown quantities $a_{x,y}^{c,s}$, $c_{x,y}^{0,1}$, and $d_{x,y}^{c,s}$ from the IERS [1] data via least-squares fitting [13, 14]. These coefficients are uniquely determined in terms of the unknown parameters in (12) and (16). We must take account of the relations

$$a_x^{c,s} \approx a_y^{s,c}, \quad d_x^{c,s} \approx d_y^{s,c}, \tag{22}$$

which describe the structural properties of the model. This also means that the processes x and y are related, which affects their statistical analysis. These equations must be satisfied fairly accurately for the independent estimates of x and y (see below).

The introduced parameters can require a correction $c_{x,y} = c_{x,y}^0 + c_{x,y}^1\tau + \ldots$ for long intervals $0 \le \tau \le \Theta$, where, for example, $\Theta \sim 10–20$ yr (similarly for $a_{x,y}^{c,s}$ and $d_{x,y}^{c,s}$). To improve the interpolation of the oscillations, the secular terms must take into account the slow evolution of the parameters in (8), (10), and (12). We can, in principle, improve the theoretical model (21) by taking into account small stochastic geophysical factors. However, this does not seem expedient now, due to the low accuracy and incompleteness of the geophysical measurements.

Let us present the calculations based on the simple least-squares technique of [12], which was applied independently to $x(\tau)$ and $y(\tau)$ using the six-parameter approximations in accordance with the model (21):

$$x(\tau) = (\xi, f(\tau)), \quad y(\tau) = (\eta, f(\tau)), \tag{23}$$
$$\xi = (\xi_1, \ldots, \xi_6)^T, \quad \eta = (\eta_1, \ldots, \eta_6)^T,$$
$$f(\tau) = (1, \tau, \cos 2\pi N\tau, \sin 2\pi N\tau, \cos 2\pi\tau, \sin 2\pi\tau)^T,$$
$$N \approx 0.845–0.85.$$

We must then determine the six-dimensional vectors ξ and η. It is reasonable first to study the effectiveness of the interpolation and forecasting of the polar motions using the simple mathematical model (23), based on the daily IERS data [1].

Figure 1 presents the theoretical curves $x^*(\tau)$ and $y^*(\tau)$ for the interpolation of the daily measurements over the eight-year time interval, $0 \le \tau \le 8$, from the beginning of 1988 to the end of 1995; the points show the measurements. The mean square deviations $\sigma_x = 0.014$ and $\sigma_y = 0.017$ demonstrate the satisfactory accuracy of the model (21), (22) for the best-fit parameters ξ^* and η^*:

$$\xi^* = (-0.041, -0.0004, -0.034, 0.0194, -0.023, -0.065)^T, \tag{24}$$
$$\eta^* = (0.300, 0.005, 0.193, 0.033, -0.060, 0.020)^T.$$

Comparing the coefficients ξ_3^* and η_4^*, ξ_4^* and η_3^* determining the Chandler oscillations and also ξ_5^* and η_6^*, ξ_6^* and η_5^* corresponding to the annual

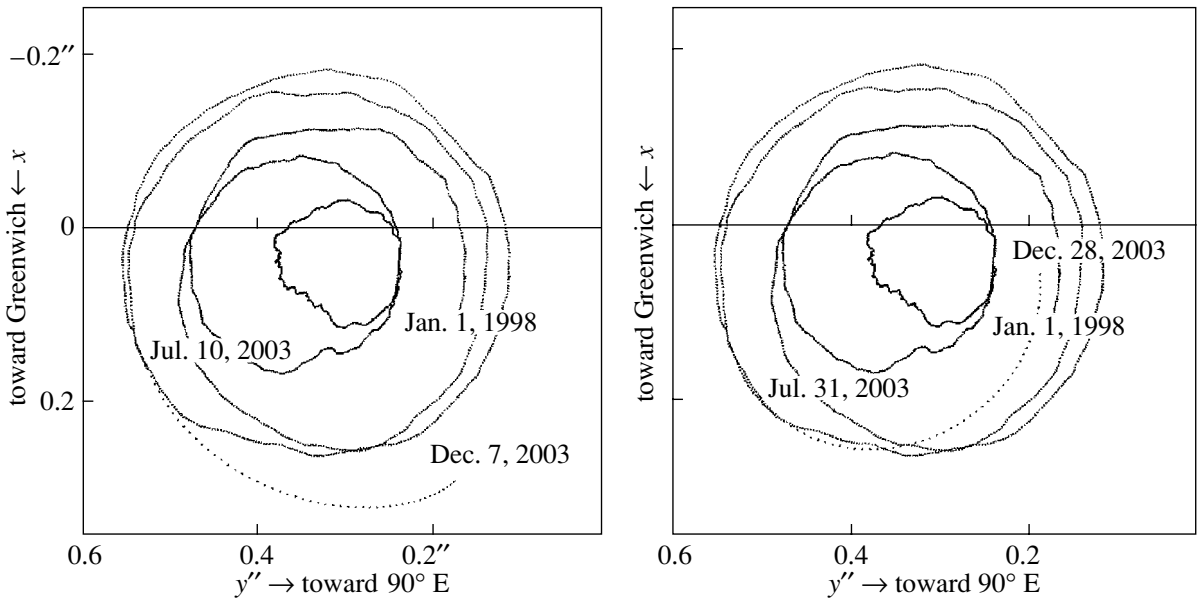

Fig. 3. Polar trajectories and their IERS forecasts (dashed curves) for July 10 and 31, 2003 (left-hand coordinate frame).

component in (24) verifies the structural properties (22) of the model (21), (23).

We have already interpolated in 2001 the observational data for the seven-year interval 1994–2000 [7]. The calculated best-fit parameters ξ^* and η^* for the model (23) and mean square deviations σ_x and σ_y are equal to

$$\begin{gathered}\xi^* = (0.039, 0.0001, 0.015, \\ 0.161, -0.046, -0.076)^T, \\ \sigma_x = 0.024; \\ \eta^* = (0.334, 0.0005, 0.162, \\ -0.0139, -0.068, 0.043)^T, \\ \sigma_y = 0.025.\end{gathered} \tag{25}$$

We proposed a two-year forecast of the polar motion for 2001–2002 using our interpolation. The dashed curve in Fig. 2 presents the theoretical curve (x^*, y^*), which is shown in two portions of one year each (for 2001 and 2002). The solid curves present the corresponding portions of the observational curve constructed using the IERS data. A comparison of the real polar trajectories and the theoretical trajectories calculated using (23) and (25) demonstrates the adequacy of the constructed model. Note the anomaly in the polar motion associated with the gravitational perturbations (the "parade of planets," see below). The fact that no geophysical treatment of the polar oscillations can explain this phenomenon was previously ignored.

The Washington Naval Observatory presents a forecast of polar motions for 100–150 days using the IERS data (IERS, EOP Product Center, http://hpiers.obspm.fr/eoppc/eop/eoppc04/eoppc04-xy.gif). The model itself and forecasting technique adopted by the IERS provide fairly inaccurate and unstable forecasts, which require corrections every week. Figure 3 demonstrates the IERS forecasts (the solid curve shows the measurements and the dots show the forecasts). The forecasts of July 10 and 31, 2003, are shown as examples; a similar situation was observed for May 18 and 27, 2004. The difference in the forecasts is compared with the maximum polar deviation. Note that the IERS data are presented in the left-hand coordinate frame.

Figure 4 presents our interpolation of the process (see the anomaly of 1999–2000) and forecast of the polar oscillations for two years (up to the end of 2005) using the eight-year daily IERS data for 1996–2003. The best-fit parameters ξ^*, η^* and mean square deviations σ_x, σ_y are equal to

$$\begin{gathered}\xi^* = (0.0314, 0.0027, -0.0494, \\ -0.0741, 0.0596, 0.1347)^T, \\ \sigma_x = 0.0175; \\ \eta^* = (0.3315, 0.001, -0.0679, \\ 0.0426, 0.1341, -0.0613)^T, \\ \sigma_y = 0.0175.\end{gathered} \tag{26}$$

A comparison of the components of the vectors ξ^* and η^* (26) verifies the validity of the approximation equations describing the properties of the model.

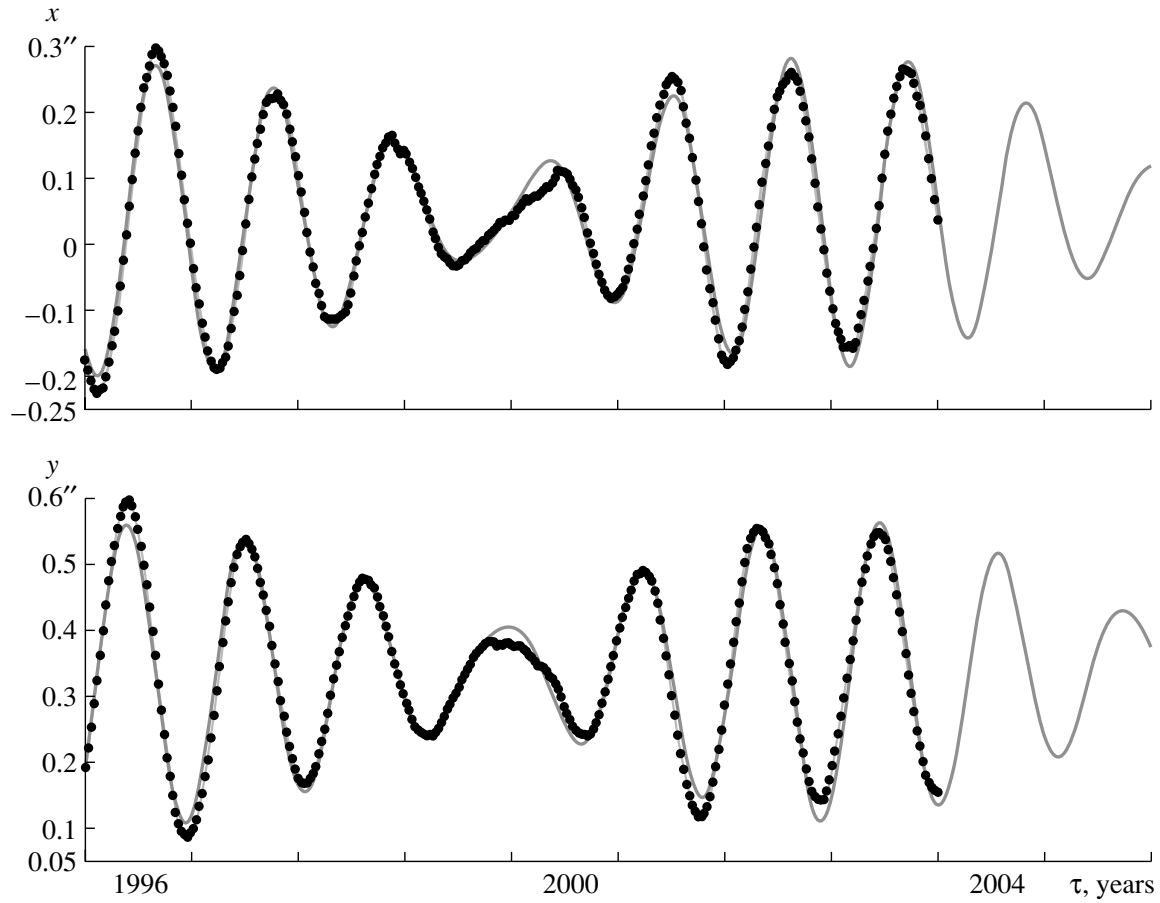

Fig. 4. The eight-year interpolation for 1996–2003 (the points show the IERS data) and two-year forecast for 2004–2005.

7. CONCLUSIONS

(1) We have studied the perturbed rotational–oscillatory motions of the Earth's pole using a model for a nearly axisymmetric, viscoelastic, rigid body. We have established that the Chandler and annual components of the oscillations are celestial-mechanical phenomena and are due to the gravitational tidal effects exerted by the Sun and Moon.

(2) We have constructed a mathematical model containing a small number of fitted parameters; the daily IERS measurements can be used to determine these parameters and verify the structural properties of the model.

(3) A statistically valid interpolation of the oscillations is provided for various time intervals. We have presented a two-year forecast that is currently being verified by the observations.

(4) We have proposed a qualitative explanation for the anomaly of the polar oscillations in 1999–2000, associated with the gravitational perturbations due to the "parade of planets." Our modeling of the polar oscillations is compared with the forecasting of the IERS.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 04-02-16303, 02-01-00252, and 02-01-00157), the program of support for Leading Scientific Schools in Russia (project no. 1627.2003.1), and the project no. MK-2093.2003.01.

REFERENCES

1. IERS Annual Reports. 1990 July 1991 bis 1999 July 2000. Central Bureau of IERS. Observatoire de Paris. 2000 July 2001 bis 2002 July 2003. Verlag BKG Frankfurt am Mein.
2. W. Munk and G. Macdonald, *The Rotation of the Earth* (Cambridge Univ. Press, Cambridge, 1960; Mir, Moscow, 1964).
3. P. Melchior, *Physique et dynamique planetaires* (Vander, Bruxelles, 1971; Mir, Moscow, 1975), Part 2.
4. H. Moritz and I. I. Mueller, *Earth Rotation: Theory and Observation* (Ungar, New York, 1987; Naukova Dumka, Kiev, 1992).
5. Yu. N. Avsyuk, *Tidal Forces and Natural Processes* (OIFZ RAN, Moscow, 1996) [in Russian].

6. N. S. Sidorenkov, *Physics of the Earth's Rotational Instabilities* (Nauka, Moscow, 2002) [in Russian].
7. V. V. Nesterov, *Standards for Basic Calculations in Astronomy. Basic Algorithms for Satellite Geodynamics* (Yanus-K, Moscow, 2001) [in Russian].
8. A. Yu. Ishlinskii, *Orientation, Gyroscopes, and Inertial Navigation* (Nauka, Moscow, 1976) [in Russian].
9. L. D. Akulenko, S. A. Kumakshev, Yu. G. Markov, and L. V. Rykhlova, Astron. Zh. **79**, 81 (2002) [Astron. Rep. **46**, 74 (2002)].
10. L. D. Akulenko, S. A. Kumakshev, Yu. G. Markov, and L. V. Rykhlova, Astron. Zh. **79**, 952 (2002) [Astron. Rep. **46**, 858 (2002)].
11. L. D. Akulenko, S. A. Kumakshev, and Yu. G. Markov, Dokl. Akad. Nauk **382** (2), 199 (2002) [Dokl. Phys. **47**, 78 (2002)].
12. V. S. Gubanov, *Generalized Least-Squares Method. Theory and Application in Astronomy* (Nauka, St. Petersburg, 1997) [in Russian].

Translated by V. Badin