Astronomy Reports, Vol. 46, No. 1, 2002, pp. 1–15. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 3–18.
Original Russian Text Copyright © 2002 by V. Polyachenko, E. Polyachenko.

Nonresonance Spiral Responses in Disk Galaxies

V. L. Polyachenko and E. V. Polyachenko

Institute of Astronomy, Russian Academy of Sciences, Pyatnitskaya ul. 48, Moscow, 109017 Russia

Received April 20, 2001

Abstract—The behavior of the gravitational potential outside the region where the main spiral arms of galaxies are located is investigated. The characteristic features of this behavior include nearly circular extensions of the main arms, which typically have an angular extent of 90°. It is natural to interpret these quarter-turn spirals as the response of the galactic disk to the gravitational potential of the main spiral arms. The theoretical models are supported by observational data for the brightness distributions in both normal (NGC 3631) and barred (NGC 1365) galaxies. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Thus far, spiral perturbations in galaxies have been studied primarily using a self-consistent approach. First, the spiral response σ_1 to a specified spiral potential Φ_1 is determined from the solution of the kinetic equation for the stars or the hydrodynamic equations for the gas; further, it is assumed that this potential itself is created by the density perturbation σ_1 (introducing the self-consistency). Precisely this formulation of the problem is used, for example, in the theory of galactic spiral density waves. In particular, the dispersion relation of Lin and Shu [1] for tightly-wound spirals was obtained in this way. Of course, the perturbations of the potential and density are determined in the same region.

We will investigate the responses of disks to the main spiral arms outside the regions where they are located. Generally speaking, there is the obvious possibility that the stellar spirals end at the edge of the stellar disk; in this case, the gaseous disk can extend appreciably further, and the gaseous spirals can likewise extend to some distance from the edge of the stellar disk. However, it is usually assumed that the strong stellar spiral arms end near a resonance (while the gaseous spirals can extend somewhat beyond it). Usually, the inner boundary of the spiral pattern is taken to be the inner Lindblad resonance (if it exists; otherwise, it is expected that the spiral structure should extend to the very center of the galaxy). Various authors have suggested the corotation radius, outer Lindblad resonance (see, for example, [2, 3]), or even the 4 : 1 resonance (e.g. [4]) as the outer boundary of the spiral pattern.

At least in some cases, the ends of the spirals can lie far from the main resonances. For example, in the galaxy NGC 3631 considered below, the outer boundary of the main spirals is outside the corotation radius but far from the outer Lindblad resonance.

Our aim in the current paper is to identify nonresonance effects associated with the ends of strong spiral arms. These are associated with a restructuring of the gravitational potential, which should obviously approach a quadrupole asymptotic far from the spirals in the case of the two-arm symmetry that is of interest to us. This restructuring occurs near the ends of the main spirals (beginning in the region they occupy) and is so rapid that it may be complete before the nearest resonance is reached, even when this resonance is located fairly close to the ends of the spirals. Note also that the resonances themselves are probably saturated in many cases, so they do not play a significant dynamical role. For these reasons, the nonresonance phenomena we consider below should be widespread in spiral galaxies.

We discuss the behavior of the potential beyond the region of the main spiral arms in detail in Section 2. Subsequently, we consider the peculiarities of the responses of the disk to the spiral-wave potential (i.e., perturbations of its surface density) separately for the outer (Section 3) and inner (Section 4) regions. The most interesting regions of the disk are those in the immediate vicinity of the main spirals. Here, characteristic quarter-turn spirals arise in relatively narrow rings: typically, the disk response forms a nearly circular arm covering about 90°. This formation of quarter-turn spirals provides a natural mechanism for the sequential (generally speaking, repeatable) extention of spirals in galaxies. The situation for the inner regions may be complicated by the possible presence of a bar; this is discussed in Section 4.

Section 5 is dedicated to observational tests of the theory developed in the previous sections. We have chosen NGC 1365 and NGC 3631 for this purpose. These galaxies belong to very different morphological types: NGC 1365 is an SB galaxy with a strong bar, while NGC 3631 is a normal spiral. We have chosen these two galaxies in order to underscore the universal character of the quarter-turn spiral phenomenon. The theory is in satisfactory agreement with the observations.

We present a brief discussion of our results and perspectives for future studies in the Conclusion (Section 6).

2. BEHAVIOR OF THE POTENTIAL INSIDE AND OUTSIDE THE REGIONS OF THE MAIN SPIRAL ARMS OF DISK GALAXIES

In this section, we investigate the behavior of the potential inside and outside the region where the main spiral arms are located, using simple models. We (briefly) consider the inner potential in order to compare it with the outer potential and identify characteristic differences between them.

2.1. Potential in the Region of the Spirals

Recall first and foremost that only the potential for tightly-wound (and, strictly speaking, multiturn) spirals can be analytically calculated. This is done using the asymptotic theory (see, for example, [5, 6]) based on expansions of the potential in inverse powers of the large quantity $kr \gg 1$ (k is the wave number at radius r). In the main order, this theory leads to the well known formula for the local relationship between perturbations of the surface density σ_1 and the potential Φ_1 [5]:

$$\Phi_1 = -\frac{2\pi G\sigma_1}{|k|}, \tag{1}$$

where G is the gravitational constant. According to (1), the density and potential are in antiphase so that, in the limiting case under consideration, the density maxima precisely coincide with minima of the potential.

The precise coincidence is already lost for open spirals, but, as before, the curve of loci of density maxima $\varphi = \varphi^{\sigma}_{\max}(r)$ follows the locus of potential minima $\varphi = \varphi^{\Phi}_{\min}(r)$ (in its central part, see Figs. 1a, 1b). It stands to reason that the potential generated by such spirals in the plane of the disk must be calculated using the general formula for the potential of a simple layer, which, for an individual azimuthal harmonic ($\propto \exp(im\varphi)$, φ is the azimuth) has the form

$$\Phi_1(r)e^{im\varphi} = -e^{im\varphi}G\int\int\frac{\sigma_1(r')e^{im\varphi'}r'dr'd\varphi'}{\sqrt{r^2+r'^2-2rr'\cos\varphi'}}. \tag{2}$$

This is done numerically; it is convenient to first transform (2) into the form (see, for example, [6])

$$\Phi_1(r) \tag{3}$$
$$= -G\int\limits_0^\infty \sigma_1(r')r'dr'\int\limits_0^{2\pi}\frac{e^{im\varphi}d\varphi}{\sqrt{r^2+r'^2-2rr'\cos\varphi}}$$
$$= -2\pi G\int\limits_0^\infty H_m(r,r')r'\sigma_1(r')dr',$$

where

$$H_m(r,r') = H_m(r',r) \tag{4}$$
$$= \frac{1}{2\pi}\int\limits_0^{2\pi}\frac{e^{im\varphi}d\varphi}{\sqrt{r^2+r'^2-2rr'\cos\varphi}}$$
$$= \frac{1}{r+r'}h_m(\zeta)$$

and

$$h_m(\zeta) = \frac{2}{\pi}\int\limits_0^{\pi/2}\frac{\cos 2mx}{\sqrt{1-\zeta\cos^2 x}}dx, \tag{5}$$
$$\zeta = \frac{4rr'}{(r+r')^2} \leqslant 1.$$

The function $H_m(r,r')$ in (3) is universal: it does not depend on σ_1 and can therefore be computed (and tabulated) ahead of time. Then, potentials corresponding to different σ_1 can be obtained simply by computing single integrals in r'. Here and below, we will be interested only in the case of two-arm symmetry ($m = 2$).

2.2. Potential Outside the Region of the Spirals

The outer potential from the spiral arms can be obtained only numerically, using either (3) or a multipole expansion (see below). However, it is known *a priori* that, sufficiently far from spirals with the azimuthal symmetry in which we are interested ($m = 2$), the potential must tend toward a quadrupole potential; that is,

$$\mathrm{Re}\,[\Phi_1 e^{2i\varphi}] \sim r^{-3}\cos 2(\varphi - \varphi_>) \tag{6}$$

in outer regions of the disk relative to the spirals (where $\varphi_>$ is some fixed azimuth), and

$$\mathrm{Re}\,[\Phi_1 e^{2i\varphi}] \sim r^{2}\cos 2(\varphi - \varphi_<) \tag{7}$$

in inner regions ($\varphi_<$ is a second fixed azimuth that, generally speaking, does not coincide with $\varphi_>$). Note that it is not obvious *a priori* that a region in which the potential has the asymptotic behavior (7) will be reached in the inner parts of the disk (the disk may not be large enough); however, numerical simulations indicate that the transition to asymptotic behavior occurs very rapidly (see Figs. 1, 2c).

It is simplest to substantiate formulas (6) and (7) (and simultaneously obtain values for the azimuths $\varphi_>$ and $\varphi_<$) starting with the general relation (3) between the density and potential if we use the fact that the asymptotic representation of the function $h_m(\zeta)$ as $\zeta \ll 1$ (i.e., for either $r' \ll r$ or $r' \gg r$) is

$$h_m \simeq \frac{(2m)!}{2^{4m}(m!)^2}\zeta^m. \tag{8}$$

Given that the spiral arms are located in a ring with inner radius r_1 and outer radius r_2, we will have, using (8),

$$\Phi_1(r) = -2\pi G\frac{(2m)!}{2^{4m}(m!)^2} \tag{9}$$

$$\times \int\limits_{r_1}^{r_2}\left[\frac{4rr'}{(r+r')^2}\right]^m \frac{r'\sigma_1(r')}{r+r'}dr'.$$

We specify the spiral surface density in the form

$$\mathrm{Re}\,[\sigma_1 e^{im\varphi}] = A(r)\cos[m\varphi - F(r)], \tag{10}$$

so that

$$\sigma_1(r) = A(r)e^{-iF(r)}. \tag{11}$$

We then obtain for the outer region ($r' \ll r$)

$$\mathrm{Re}\,[\Phi_1 e^{im\varphi}] \simeq A_> r^{-(m+1)}\cos m(\varphi - \varphi_>), \tag{12}$$

where

$$A_> = -2\pi G\frac{(2m)!}{2^{2m}(m!)^2}\sqrt{F^2_{c(>)} + F^2_{s(>)}} \tag{13}$$

$$\tan m\varphi_> = F_{s(>)}/F_{c(>)}, \tag{14}$$

$$F_{s(>)} = \int\limits_{r_1}^{r_2} r'^{m+1}A(r')\sin F(r')dr', \tag{15}$$

$$F_{c(>)} = \int\limits_{r_1}^{r_2} r'^{m+1}A(r')\cos F(r')dr'.$$

In the case $m = 2$, as expected, we obtained the quadrupole asymptotic (6) and showed that the azimuth $\varphi_>$ can be computed using the formula

$$\varphi_> = \frac{1}{2}\arctan\frac{\int_{r_1}^{r_2} r'^3A(r')\sin F(r')dr'}{\int_{r_1}^{r_2} r'^3A(r')\cos F(r')dr'}. \tag{16}$$

Analogously, for the inner region ($r' \gg r$),

$$\mathrm{Re}\,[\Phi_1 e^{im\varphi}] \simeq A_< r^m\cos m(\varphi - \varphi_<), \tag{17}$$

where

$$A_< = -2\pi G\frac{(2m)!}{2^{2m}(m!)^2}\sqrt{F^2_{c(<)} + F^2_{s(<)}}, \tag{18}$$

$$\tan m\varphi_< = F_{s(<)}/F_{c(<)}, \tag{19}$$

$$F_{s(<)} = \int\limits_{r_1}^{r_2}\frac{A(r')}{r'^m}\sin F(r')dr', \tag{20}$$

$$F_{c(<)} = \int\limits_{r_1}^{r_2}\frac{A(r')}{r'^m}\cos F(r')dr'.$$

In accordance with (17), when $m = 2$, we have an asymptotic of the form (7), where the azimuth $\varphi_<$ is given by the formula

$$\varphi_< = \frac{1}{2}\arctan\frac{\int_{r_1}^{r_2}\frac{A(r')}{r'^2}\sin F(r')dr'}{\int_{r_1}^{r_2}\frac{A(r')}{r'^2}\cos F(r')dr'}. \tag{21}$$

Regular series of asymptotic expansions of the spiral potential in powers of r'/r for the outer region (in powers of r/r' for the inner region) yield a multipole representation of this potential. Using the identity

$$\frac{1}{\sqrt{r^2 + r'^2 - 2rr'\cos\varphi}} \tag{22}$$

$$= \frac{1}{R_>}\sum_{n=0}^{\infty}\left(\frac{r'}{R_>}\right)^n\left(\frac{R_>}{r}\right)^{n+1}P_n(\cos\varphi),$$

for the outer region ($R_>$ is the radius of the outer edge of the spiral), we obtain for $m = 2$ from (2)

$$\Phi_1(r,\varphi) = -G\mathrm{Re}\,e^{2i\varphi}\sum_{n=2}^{\infty}\left(\frac{R_>}{r}\right)^{n+1}Q_2^{(n)}p_n, \tag{23}$$

$$Q_2^{(n)} = \int\limits_{r_1}^{r_2}\left(\frac{r'}{R_>}\right)^{n+1}A(r')e^{-iF(r')}dr', \tag{24}$$

$$p_n = \int\limits_0^{2\pi}d\varphi' e^{2i\varphi'}P_n(\cos\varphi') \tag{25}$$

$$= \left\{\frac{(2k-3)!!}{(k-1)!}\,\frac{\prod_{i=1}^{k-1}(2k+1-i)}{\prod_{i=1}^{k-1}(4k+1-2i)}\right\}\cdot\frac{(4k-1)!!}{2^{2k-1}(2k)!}\cdot\pi$$

for even n ($k = n/2$) and $p_n = 0$ for odd n, so that the sum in (23) runs only over even n. The last equality of (25) can easily be derived from the expansion of $P_n(\cos\varphi)$ in a multiple-angle cosine series in [7] (note that there are errors in the corresponding expansion presented in the handbook [8]). In particular, we find from (25) $p_2 = 3\pi/4$, $p_4 = 5\pi/16$, and $p_6 = 105\pi/2^9$. Note that, for $p_2 = 3\pi/4$, we obtain a quadrupole potential from (23), which coincides with that found

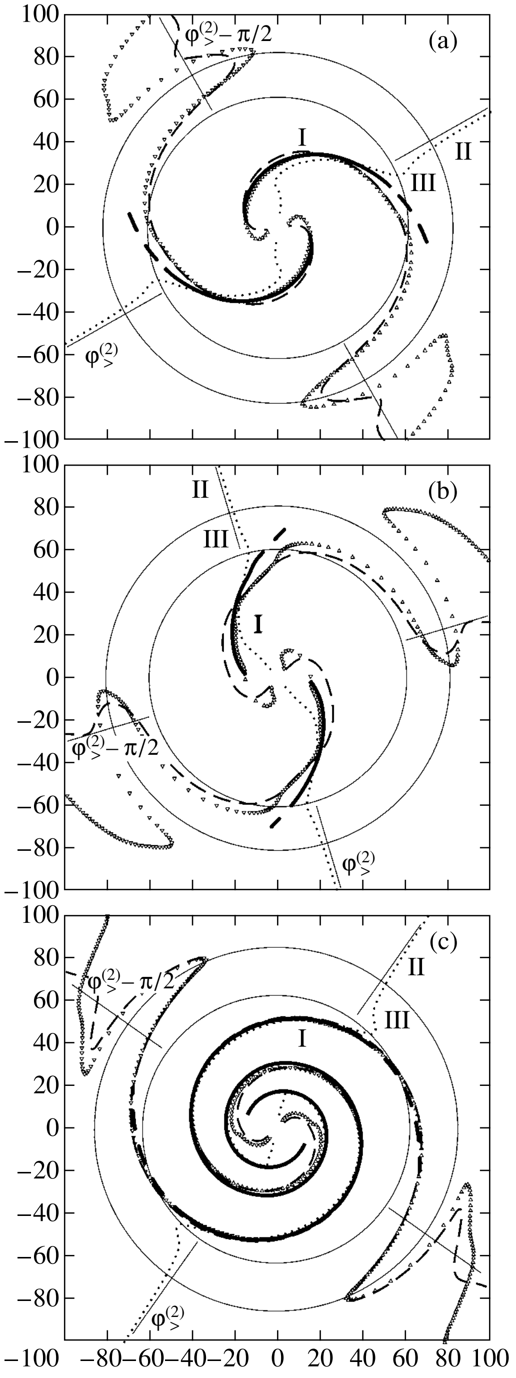

Fig. 1.

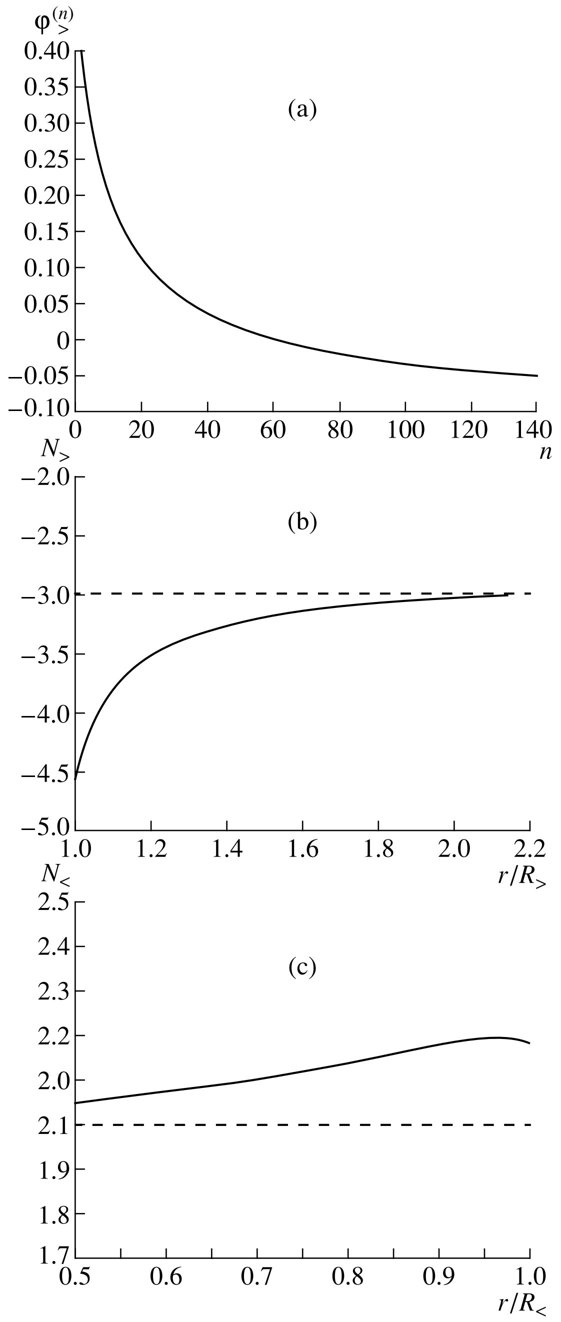

Fig. 2.

earlier in (12), (13) if we substitute $m = 2$ in these formulas.

It is straightforward to obtain the recurrence relation

$$p_{k+1} = p_k \frac{(2k-1)(2k+3)}{4k(k+2)}$$

from (25). In turn, we can now use this, for example, to derive the asymptotic behavior of the coefficients $p_n = p_{2k} \equiv q_k$ for large n or k. Since, in this case,

$$\begin{aligned} q_{k+1} - q_k &\simeq \frac{dq_k}{dk} \\ &= q_k \left\{ \frac{(2k-1)(2k+3)}{4k(k+2)} - 1 \right\} \simeq -\frac{q_k}{k}, \end{aligned}$$

it is obvious that $q_k \simeq c/k$ (or $p_n \simeq 2c/n$); the value of the constant c is determined numerically and turns out to be $c \simeq 0.637$.

Fig. 1. Loci of the potential minima (dotted curves, which asymptotically approach azimuth $\varphi_>^{(2)}$ at large r) and density maxima of the main spirals (bold curves), the second derivative of the potential (dashed curves), and the responses of the density (triangles) in disk models. The rotation curve corresponds to that of the galaxy NGC 3631 [10], with $r_{\mathrm{CR}} = 42$ and $r_{\mathrm{OLR}} = 150$ (the rotation curve is smoothed in the central region to ensure that the inner Lindblad resonances are absent). The disk is exponential, with surface density $\sigma_0(r) = \sigma_0\ \exp(-r/r_d)$ for $r > 30$, and $\sigma_0(r) = \sigma_0\ \exp(-30/r_d) = \mathrm{const}$ for $r < 30$. The main spirals are logarithmic, with $\sigma_1 = A(r)\ \cos(2\varphi + B\ln(r/b))$; the resonances are suppressed in accordance with (36) when $\gamma_0 = 0.15$. The circles separate regions with different potential behaviors: Region I exhibits spiral behavior, Region II multipole behavior, and Region III is a transition region between these (only the outer parts of Regions II and III are shown; there are also analogous inner regions). (a) $B = 4$ and $r_d = 50$; (b) $B = 2$ and $r_d = 60$; (c) $B = 12$ and $r_d = 30$ (when r_d is appreciably larger than these values, we obtain responses with lengths of $3\pi/2$, or sometimes π). The quarter-turn spirals lie in narrow rings in Region III; in the case of open spirals (a) and (b), the quarter-turn spirals are very obvious, while they are more difficult to distinguish in the case of tightly-wound spirals (c).

Fig. 2. (a) Azimuths $\varphi_>^{(n)}$ of the phases of the complex multipole moments for the logarithmic main spiral $\sigma_1 = A(r)\ \cos(2\varphi + B\ln(r/b))$ for $B = 4$. Actually for $n \to \infty$ $\varphi_>^{(n)}$ tend to the azimuth of the line passing through the outer ends of the main spiral (in this case, $\varphi_>^{(\infty)} \simeq -0.1$). (b) Effective power-law index $N_> = d\log\bar{\Phi}_>/d\log r$ ($\bar{\Phi}_>$ is the amplitude of the potential in the outer region) as a function of $r/R_>$ ($R_> = 70$ is the radius of the outer ends of the main spirals); $N_> \to -3$ for large $r/R_>$ (as expected in the quadrupole limit). (c) Effective power-law index $N_< = d\log\bar{\Phi}_</d\log r$ ($\bar{\Phi}_<$ is the amplitude of the potential in the inner region) as a function of $r/R_<$ ($R_< = 15$ is the radius of the inner rings of the main spirals); $N_< \to 2$ for small $r/R_<$ (the quadrupole limit).

If we retain two terms (with $n = 2$ and $n = 4$) in (23),

$$\Phi_1(r,\varphi) = -\frac{\pi G}{16}\left(\frac{R_>}{r}\right)^3 \qquad (26)$$
$$\times\ \mathrm{Re}\, e^{2i\varphi}\left(12Q_2^{(2)} + 5Q_2^{(4)}\left(\frac{R_>}{r}\right)^2\right),$$

we can obtain an r-dependent correction to the azimuth $\varphi_a = \varphi_>$ of the asymptotic to which the curve $\varphi = \varphi_{\min}^{\Phi}(r)$ tends for large r [see (12), (16)]. It is easy to show that

$$\Phi_1 \propto \cos[2(\varphi - \varphi_0(r))], \qquad (27)$$

where

$$\tan 2\varphi_0(r) = \frac{q_{2i} + 5q_{4i}R_>^2/12r^2}{q_{2r} + 5q_{4r}R_>^2/12r^2}, \qquad (28)$$

and in (28) $Q_2^{(2)} = q_{2r} - iq_{2i}$ and $Q_2^{(4)} = q_{4r} - iq_{4i}$. Denoting $\varphi_0 = \varphi_a + \delta\varphi(r), where(\tan 2\varphi_a = q_{2i}/q_{2r})$, we find from (28)

$$\delta\varphi(r) \simeq \frac{5R_>^2}{4r^2}\frac{q_{4i}q_{2r} - q_{2i}q_{4r}}{q_{2r}^2 + q_{2i}^2}. \qquad (29)$$

The expansion for the inner region that is analogous to (23) has the form

$$\Phi_1(r,\varphi) = -G\mathrm{Re}\, e^{2i\varphi}\sum_{n=2}^{\infty}\left(\frac{r}{R_<}\right)^n Q_1^{(n)}p_n, \qquad (30)$$

where

$$Q_1^{(n)} = \int_{r_1}^{r_2}\left(\frac{R_<}{r'}\right)^n A(r')e^{-iF(r')}dr', \qquad (31)$$

and $R_<$ is the radius of the inner edge of the main spiral.

The potential in the immediate vicinity of the spirals can only be obtained numerically, using formulas (3)–(5) or (23), (30). We present a typical example of such computations in Fig. 1a. This figure corresponds to a simple model with a two-arm logarithmic spiral

$$\sigma_1 = A(r)\cos(2\varphi - F(r)), \qquad (32)$$
$$F(r) = -B\ln(r/b),$$

specified in the ring $15 < r < 70$. The amplitude $A(r)$ is taken to be constant ($A = 7$) in the region $30 < r < 55$, and the edges are smoothed using a $\cos^2 x$ function (note that the modeling results change little qualitatively if other representations for σ_1 are used; for example, we experimented with using segments of two exponents, growing in the central region and decaying on the periphery, joined by a spline fit). The parameters $b = 0.6$ and $B = 4$; this value of B corresponds to a fairly open spiral.

As expected, this figure clearly shows the presence of well-defined regions with fundamentally different behaviors for the potential. Region I exhibits spiral behavior and coincides with a large portion of the region where the spiral arms are located; here, the curve of loci of the potential minima $\varphi = \varphi_{\min}^{\Phi}(r)$ roughly follows the curve of locus of the surface-density maxima $\varphi = \varphi_{\max}^{\sigma}(r)$.

Region II is made up of two regions (outer and inner) exhibiting multipole behavior. The outer Region II exhibits a quadrupole potential at large r; here, the curve $\varphi = \varphi^{\Phi}_{\min}(r)$ degenerates into the radial line $\varphi = \varphi_a$.

Regions I and II are separated by a narrow transition Region III, in which there is a rapid transformation of the potential from one type to the other. As we will see in the following Section, the formation of characteristic nearly-circular spirals extending the main arms is associated precisely with the transformation of the potential in Region III.

The most interesting thing here is that, after leaving the spiral, the curve $\varphi = \varphi^{\Phi}_{\min}(r)$ very rapidly and without any significant deviations (even initial deviations) approaches the asymptotic quadrupole regime $\varphi = \varphi_a$. This occurs long before this curve enters the wave zone ($r/r' \gg 1$), where quadrupole behavior is natural. This phenomenon has its origin in a synchronization of the phases $\varphi^{(n)}_>$ of the complex multipole moments

$$\sim Q_2^{(n)} = q_r^{(n)} - iq_i^{(n)}$$
$$= \sqrt{(q_r^{(n)})^2 + (q_i^{(n)})^2}\, e^{-2i\varphi^{(n)}_>} (\tan 2\varphi^{(n)}_> = q_i^{(n)}/q_r^{(n)}),$$

and precisely the angles $\varphi^{(n)}_>$ are close for all n. The reason is that, at large (and even not very large) n, the main contribution to the integral (24) determining $Q_2^{(n)}$ is made by the maximum r', i.e., by regions near the spiral ends. Thus, the direction of the line connecting the outer ends of the main spirals corresponds to the azimuth $\varphi^{(\infty)}_>$. The remaining $\varphi^{(n)}_>$ (for $n = 4, 6, \ldots$) lie between $\varphi^{(2)}_>$ and $\varphi^{(\infty)}_>$; as n increases, the azimuth $\varphi^{(n)}_>$ monotonically approaches $\varphi^{(\infty)}_>$, which serves as a convergence point for the sequence of $\varphi^{(n)}_>$. Direct computation of the angles $\varphi^{(n)}_>$ using (16) confirms the validity of these qualitative arguments (Fig. 2a).

It thus follows that, in any case, the deviation of the curve $\varphi = \varphi^{\Phi}_{\min}(r)$ from the asymptotic quadrupole direction (after leaving the spiral) cannot exceed $\Delta\varphi = \varphi^{(2)}_> - \varphi^{(\infty)}_>$. Actually, this curve turns toward its asymptotic position along the azimuth $\varphi^{(2)}_>$ not at the end of the spiral (in the azimuth), but appreciably earlier. This is because the contributions of multipole orders to the potential fall off due to both the corresponding powers $(r'/r)^{(n+1)}$ (for $(r'/r) < 1$) and the decrease of the coefficients p_n as n increases (as we saw, asymptotically when $n \gg 1$, $p_n \sim 1/n$). The decrease in the amplitude $A(r)$ at the edge of the spiral also plays a role. Finally, we note that very open spirals themselves can lie within a fairly narrow azimuth interval.

The situation in the inner region is analogous in many respects. In particular, here also, the curve $\varphi = \varphi^{\Phi}_{\min}(r)$ very rapidly and without appreciable deviations approaches the asymptotic direction $\varphi^{(2)}_<$ after leaving the spiral. In this case also, the synchronization of the phases $\varphi^{(n)}_<$ of the complex multipole moments

$$\sim Q_1^{(n)} = p_r^{(n)} - ip_i^{(n)} = \sqrt{(p_r^{(n)})^2 + (p_i^{(n)})^2}\, e^{-2i\varphi^{(n)}_<}$$
$$\times (\tan 2\varphi^{(n)}_< = p_i^{(n)}/p_r^{(n)})$$

can be explained by the fact that, at large n, the main contributions to the integral (31) are from the spiral ends (which now correspond to the minimum radii r'). Therefore, (i) the azimuth $\varphi^{(\infty)}_<$ is determined by the direction of the line connecting the inner ends of the main spirals; (ii) the sequence $\varphi^{(n)}_<$ is monotonic, and $\varphi^{(\infty)}_<$ is a convergence point for this sequence; and (iii) in principle, the curve $\varphi = \varphi^{\Phi}_{\min}(r)$ cannot go beyond the edge of the main spirals, and in reality the maximum deviation of this curve from the direction $\varphi^{(2)}_<$ is much smaller (for the same reasons indicated above for the outer region).

In light of the above discussion, the potential in the outer region can be written accurately in the form

$$\Phi_1(r,\varphi) \simeq -\bar{\Phi}_>(r)\cos[2(\varphi - \varphi^{(2)}_>)], \qquad (33)$$

while the potential in the inner region can be written in the form

$$\Phi_1(r,\varphi) \simeq -\bar{\Phi}_<(r)\cos[2(\varphi - \varphi^{(2)}_<)], \qquad (34)$$

where $\bar{\Phi}_>(r)$ and $\bar{\Phi}_<(r)$ are positive real functions having a quadrupole asymptotics sufficiently far from the spirals: $\bar{\Phi}_>(r) \to \mathrm{const}_1 \times r^{-3}$ and $\bar{\Phi}_<(r) \to \mathrm{const}_2 \times r^2$, with both constants positive.

The rapidity with which the full asymptotic is reached (both in terms of the direction $\varphi^{(2)}_>$ and in terms of the radius dependence of the amplitude $\bar{\Phi}_>(r)$) is illustrated by Fig. 2b, which presents the dependence of the effective power-law index $N_> = d\ln\bar{\Phi}_>/d\ln r$ on $r/R_>$ for a typical model described above, corresponding to Fig. 1a. We can see that the power-law index $N_>$ rapidly approaches its asymptotic value $N_> = -3$.

For the inner region, the effective power-law index is $N_< = d\ln\bar{\Phi}_</d\ln r$; Fig. 2c presents the dependence $N_< = N_<(r/R_<)$. Values close to the asymptotic value $N_< = 2$ are reached very rapidly in this region.

Let us briefly consider how the potential is modified when the parameters of the main spiral are changed. First and foremost, it is obvious that, when the size of the region of density smoothing at the spiral ends is increased, the asymptotic direction $\varphi^{(2)}_>$

should shift inward somewhat, while $\varphi_<^{(2)}$ should shift outward. It stands to reason that there should not be any variations in the curve $\varphi = \varphi_{\min}^{\Phi}(r)$ when the amplitude $A(r)$ is varied (in view of the linearity of the Poisson equation). It turns out that the appearance of this curve does not depend appreciably on the degree of winding of the spirals. When the radius $R_<$ of the inner edge of a fairly open spiral is increased, the azimuth $\varphi_>$ is shifted toward the outer edge.

To conclude this section, let us say a few words about the influence of a bar on the total potential, equal to the sum of the potentials of the spirals and bar. In principle, the bar potential could be so strong that Region I, with spiral potential behavior, is completely absent from the galaxy. However, we are not interested in such cases here and consider only situations when the spiral is sufficiently powerful for the spiral potential to exceed the bar potential, at least over some portion of the spiral length (of course, not necessarily directly from the end of the bar). It is clear that this imposes specific constraints on the relative masses of the bar and spirals. A separate study is required to determine these criteria.

The general character of the behavior of the total potential in the region of the galactic disk outside the spirals remains the same as in the absence of a bar. In particular, this is true of the form of the characteristic curve $\varphi = \varphi_{\min}^{\Phi}(r)$. Therefore, we will not repeat here the results of computing the potentials for barless spirals (described above).

We will consider the region of the disk inside the spirals only in cases when there is some gap between the end of the bar and the inner edge of the spiral (if there is no gap, we enter the bar immediately after leaving the spiral). Depending on the orientation of the bar relative to the inner ends of the spirals, a large variety of configurations for the total potential and the curve $\varphi = \varphi_{\min}^{\Phi}(r)$ are possible. We will present some of these at the end of the next Section (together with the responses to these potentials).

3. RESPONSES OF DISKS IN THEIR OUTER REGIONS

We now turn to an investigation of disk responses that can be observed directly as (generally speaking, weaker) extensions of the main spiral arms.

Let us assume that all the nonaxially symmetric structure—which, in the case of an SB galaxy, for example, includes the bar and all the spiral arms (both the main arms and their extensions)—primarily corresponds to some normal disk mode with two-arm azimuthal symmetry. The main characteristic of the mode is that all its components have a single angular velocity Ω_p (or angular frequency $\omega = 2\Omega_p$).

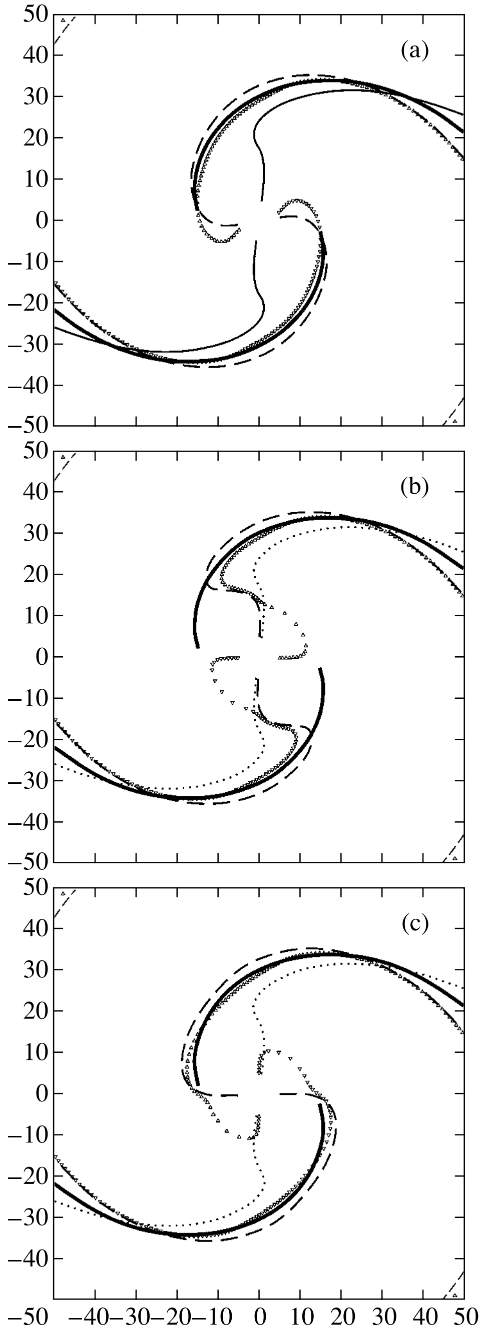

Fig. 3. Responses of the inner regions of disks under the joint action of the main spirals and the bar. The spiral is assumed to be logarithmic, $\sigma_1 = A(r)\cos(2\varphi + B\ln(r/b))$, with $B = 4$. The inner Lindblad resonances are absent. (a) Same as in the inner region in Fig. 1a (with no bar); (b) with a horizontal bar; (c) with a vertical bar.

Although they represent a component of the total normal mode, the characteristic extensions of the main spiral arms are nevertheless structures induced by another component (the main spirals).

To describe the response of a galactic disk to the potential of a spiral (and bar), we adopt a disk model with circular orbits. It is clear that this approximation is fully justified when the velocity perturbations due to the non-axially symmetric potential exceed the thermal velocities in the stellar subsystem considered. There is no doubt that this approximation is suitable for gaseous clouds throughout the galaxy, and also for stars in the distant regions of the disk outside the spirals, while it may represent only a crude approximation for stars in the inner region of the disk. However, we will see that even this simple model leads to results that are in reasonably good qualitative agreement with observational data for both outer and inner regions.

On the other hand, we have to use a cool-disk approximation, since the required analytical expression for the disk response to an arbitrary potential, including the potential of an open spiral or bar, can be derived only in a cool-disk model. We are not yet able to calculate the disk response to this type of potential, even if the peculiar velocities of the stars are small (but nonzero). Both of the galaxies we consider in Section 5 (NGC 1365 and NGC 3631) have very open main spirals. Therefore, a tightly-wound spiral approximation (which is the only one for which a general expression for the response of a disk with nearly circular orbits is known [1]) is not appropriate, even in the region where the spirals are located; it is even less suitable in regions beyond the spirals, where, as we saw in the previous Section, the potential rapidly approaches a multipole regime. Note also that the approach to a multipole regime is no less rapid for tightly wound main spirals. Therefore, in this case also, studies of disk responses to a potential outside the spirals can be conducted only numerically or using a cool-disk model.

We will restrict our treatment to a linear analysis, assuming that all perturbations of various quantities are proportional to $\exp(-i\omega t + im\varphi)$, where t is time, $m = 2$, and $\omega = 2\Omega_p + i\gamma$ is the complex frequency, comprised of the rotational velocity Ω_p and the growth rate γ. Linearizing the Euler hydrodynamical equations with zero pressure, we can express the perturbed velocities v_r and v_φ in terms of the perturbed gravitational potential Φ_1. Substituting these expressions for v_r and v_φ into the linearized continuity equation, we obtain a relation between the perturbations of the surface density $\sigma_1(r)$ and the potential $\Phi_1(r)$:

$$\sigma_1(r) = -\frac{1}{r}\frac{d}{dr}\left(r\varepsilon\frac{d\Phi_1}{dr}\right) + \frac{4}{r^2}\varepsilon\Phi_1 + \frac{4}{r\omega_*}\Phi_1\frac{d}{dr}(\varepsilon\Omega). \tag{35}$$

Here, $\varepsilon(r) = \sigma_0(r)/(\omega_*^2 - \kappa^2)$ is the gravitational analog of the dielectric constant, $\sigma_0(r)$ is the unperturbed density, $\kappa(r)$ is the epicyclic frequency, $\kappa^2 = 4\Omega^2 + rd\Omega^2/dr$, $\Omega(r)$ is the local angular velocity of the galactic disk, and $\omega_*(r) = \omega - 2\Omega(r)$. Formula (35) was first presented in this form in [9]. By definition, the real quantity $\mathrm{Re}\,[\sigma_1(r)\exp(2i\varphi)] = A(r)\cos[2\varphi - F(r)]$ represents the response of the density (or simply the "response") to the potential Φ_1; $A(r) > 0$ is the real amplitude, and $F(r) = \int^r k(r')dr'$ is the phase of the spiral wave (k is the wave number). Similarly, we can write for the potential $\Phi_2(r,\varphi) = \mathrm{Re}[\Phi(r)\exp(2i\varphi)] = A_\Phi(r)\cos[2\varphi - F_\Phi(r)]$, $F_\Phi(r) = \int^r k_\Phi(r')dr'$ (generally speaking, $k(r)$ and $k_\Phi(r)$ do not coincide).

In view of the rapid radial variations of the potential, both in the main spiral and in the transition from the spiral to the multipole regime, we usually have $\sigma_1(r) \cong -\varepsilon d^2\Phi_1/dr^2$ in place of (35). This can also be shown via direct computation and comparison of terms proportional to Φ_1'', Φ_1', and Φ_1 in (35); as a rule, the first of these is appreciably larger than the others. This is also confirmed by the curves in Fig. 1, which show the results of using (35) to compute the loci of maxima (at each radius) of the response to the potential of a logarithmic main spiral. Computation of the response using (35) requires that we adopt some rotation law for the disk, $\Omega = \Omega(r)$. The appearance of the response depends critically only on the sign of ε. As a rule, the ends of the spirals lie in the region where $\omega_*^2 < \kappa^2$; i.e., between the Lindblad resonances [1]. In this case, $\varepsilon < 0$, and the response $\sigma_1(r)$ varies in the same way as the second derivative of the potential, $\sigma_1(r) \sim \Phi_1''$. This most typical case is presented in Fig. 1 (see the figure caption for a description of the specific model adopted).

Note that, in the region of spiral behavior of the potential, $\Phi_1'' \sim -k_\Phi^2\Phi_1$, while, in the multipole regime, $\Phi_1'' \sim +n(n+1)\Phi_1$. Hence, we can see that, if we write $\sigma_1(r) = B(r)\Phi_1(r)$, the phase of the complex function $B(r)$ changes by π in the transition from the spiral beginning of the response to its multipole end. This corresponds to the length of the two-arm response in azimuth, equal to $\pi/2$ (since variations of the phase ΔF and the turn angle of the two-arm spiral $\Delta\varphi$ are related by the expression $\Delta F = 2\Delta\varphi$, and, in this case $\Delta F = \pi$).

In addition to the fact that the disk response closely follows the relation $\sigma_1(r) \sim \Phi_1''$, we will make note of certain other characteristics of typical responses. Most importantly, we can see that, before the quarter-turn spiral becomes the asymptotic radial line $\varphi = \varphi_> \pm \pi/2$ (Fig. 1), it experiences several decaying oscillations about the asymptotic azimuth (recall that the curve $\varphi = \varphi_{\min}^{\Phi}(r)$ approaches the asymptotic without such oscillations). In addition,

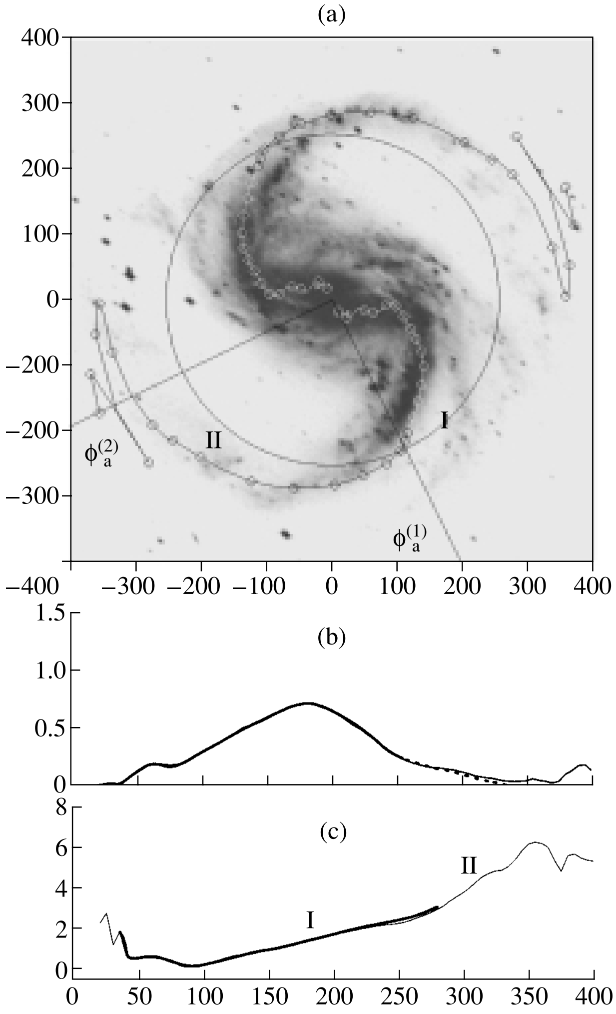

Fig. 4.

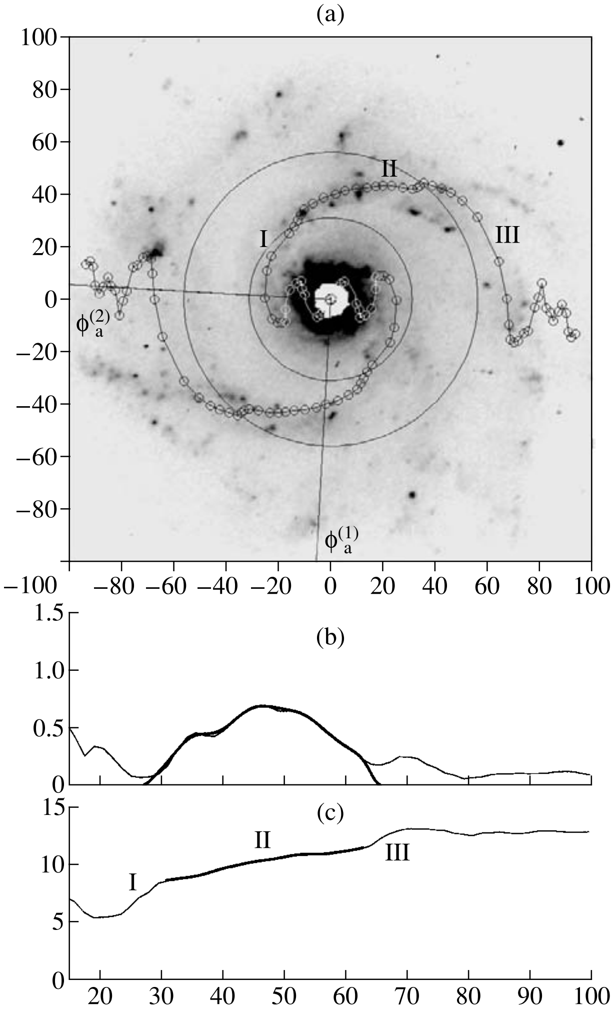

Fig. 5.

Fig. 4. (a) Deprojected image of the galaxy NGC 1365 with superposed Fourier harmonic $m = 2$ of the B brightness distribution and the (b) amplitude $A(r)$ and (c) phase $F(r)$ of this Fourier harmonic. The image was obtained from the NED archive. I is the region of the main spirals, and II that of the quarter-turn spirals (these two regions are separated by a circle; the meaning of Regions I and II here differs from that in Fig. 1). The bold curves show the smoothed functions used to compute the response of the galactic disk (see Fig. 6).

Fig. 5. (a) Deprojected R image of the galaxy NGC 3631 and the (b) amplitude $A(r)$ and (c) phase $F(r)$ of the corresponding Fourier harmonic. I is the region of the inner quarter-turn spiral; II the region of the main spiral; and III the region of the outer quarter-turn spiral (Regions I, II, and III are separated by circles; the meaning of these regions differs from that in Fig. 1). The bold curves show the smoothed functions used to compute the galactic-disk response (Fig. 7). The image of the galaxy was obtained from the ING archive. The quarter-turn spirals in different wavelength ranges (including in the $\lambda = 21$ cm radio line) are similar. The possible presence of a minibar in the center does not affect the formation of the quarter-turn spirals, although the bar can strongly influence the shape of the multipole oscillation closer to the center. This is also true with regard to the effect of inner unsaturated resonances.

the deviations of the curve of loci of the response maxima $\varphi = \varphi^{\sigma}_{\max}(r)$ from its asymptotic behavior (i.e., the amplitudes of these oscillations) are much larger than the corresponding deviations for the curve $\varphi = \varphi^{\Phi}_{\min}(r)$. Note also that these oscillations are amplified in the transition from open to tightly-wound spirals.

The beginning of the quarter-turn spirals (and simultaneously the end of the main spirals) can be defined using several slightly different methods. The most natural of these is to associate their beginning with a sharp decrease in the inclination i of the spiral response [for the logarithmic main spiral (32), the inclination is constant, $\tan i = 2/B$]. A sharp change in the inclination can be seen even in the response for

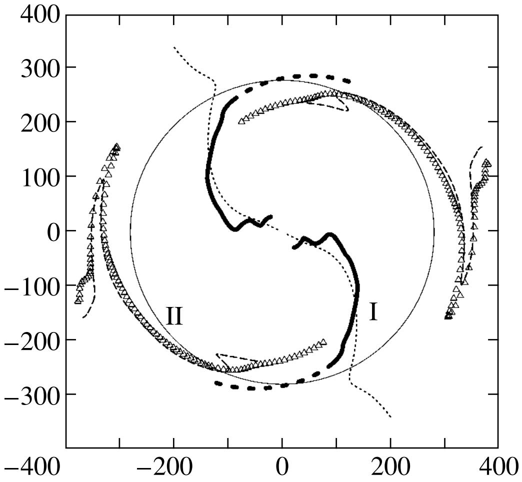

Fig. 6. Loci of maxima of the density response at each radius (triangles, II) to the gravitational potential of the main spirals (bold solid curves, I) for the galactic disk of NGC 1365; the regions of the main spirals and quarter-turn spirals are separated by a circle. Dotted curves are used for the loci of the minima of the two-arm potential $\Phi_2(r, \varphi)$. The dashed curves show the loci of the maxima of the function $\mathrm{Re}\,[\Phi'' \exp(2i\varphi)]$. Qualitatively, the appearance of the quarter-turn spirals depends only weakly on the specific values of the disk and spiral-wave parameters. As we can see, the curve of loci of the response maxima and the quarter-turn spirals in Fig. 4a are very similar.

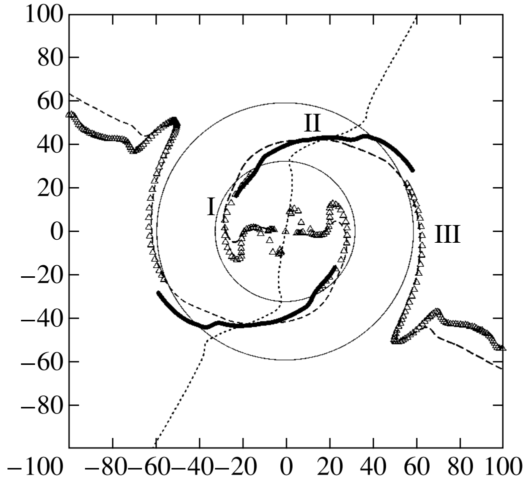

Fig. 7. Loci of maxima of the density response at each radius (triangles, I and III) to the gravitational potential of the main spirals (bold solid curves, II) for NGC 3631. The remaining notation is the same as in Fig. 6. Qualitatively, the appearance of the quarter-turn spirals depends only weakly on the specific values of the disk and spiral-wave parameters. As we can see, the curve of loci of the response maxima and the quarter-turn spirals in Fig. 5a are very similar.

a tightly-wound spiral (Fig. 1c) and is very distinct in the responses for open spirals (Figs. 1a, 1b). The degree of winding of the quarter-turn spirals themselves is not very dependent on the inclination of the main spirals.

Another way to define the beginning of the quarter-turn spirals is to associate this with the place where the total potential acquires a multipole form; the difference of this from the first definition is only slight. Recall that the angle between the azimuths of the quadrupole asymptotic of the potential and the disk response is precisely 90° (at sufficiently large or small r).

In the Conclusion (see Fig. 8), we will briefly discuss the peculiarities of the disk response in those rare, possibly even hypothetical, cases when the radius of the outer Lindblad resonance r_{OLR} is smaller than the radius of the outer edge of the spiral $R_>$ (or more precisely, smaller than the radius of the point at which the transition from the spiral to the multipole potential regime begins).

4. RESPONSES OF DISKS IN THEIR INNER REGIONS

The responses of disks in central regions not containing a bar are presented in Fig. 1, as are the responses in the outer regions. Most often, they extend the main spirals inward in the disk by roughly a quarter-turn turn. We saw above that this occurs when, first, the main spiral ends before reaching the inner Lindblad resonance and, second, the term $\sim \Phi''$ exceeds the remaining terms ($\sim \Phi'$ and $\sim \Phi$) in expression (35) for the response. Note that, in the inner regions, this last condition is usually satisfied, but with less of a margin than in the outer region.

The most important difference from the outer region is associated with the possible influence of the bar on the potential and the response of the disk in the gap between the ends of the bar and the inner edges of the main spiral. Figure 3 shows some configurations corresponding to different ratios of the masses in the spirals and bar (and with different mutual orientations). We can see that the responses are indeed more varied in this case. The influence of the inner Lindblad resonances (if present) on the disk response in the central region can be appreciable (see point 4 in the Conclusion).

5. EXAMPLES OF QUARTER-TURN SPIRALS IN REAL GALAXIES

Figure 4a presents a deprojected image of the well-studied SB galaxy NGC 1365 superposed with the Fourier harmonic of the brightness distribution corresponding to the dominating two-arm symmetry. The surface density σ_2, which is proportional to the

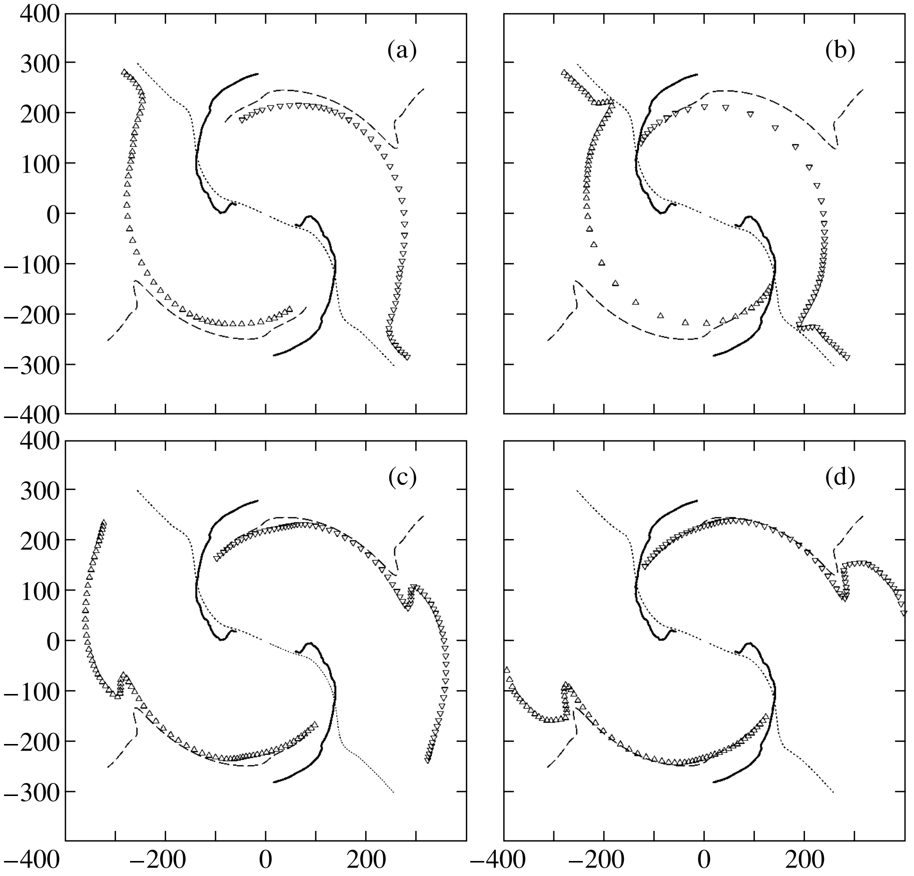

Fig. 8. Outer responses of disks (triangles) induced by the main spirals of NGC 1365 in models with the same rotation curve (from [13]) but different corotation radii. In (a) and (b), $r_{\rm CR} = 145''$ ($r_{\rm OLR} = 215''$); this value of $r_{\rm CR}$ is proposed in [13]. In (a), the resonances are suppressed in accordance with (36); (b) presents the purely resonance response for $\gamma = \gamma_0 = \text{const}$. In (c), $r_{\rm CR} = 180''$ ($r_{\rm OLR} = 280''$), and in (d), $r_{\rm CR} = 200''$ ($r_{\rm OLR} = 325''$). In (c) and (d), the resonances are suppressed in accordance with (36).

brightness, can be written in the polar coordinates (r, φ) in the form $\sigma_2(r,\varphi) = \mathrm{Re}\ [\sigma(r)\exp(2i\varphi)] = A(r)\cos[2\varphi - F(r)]$, where the complex amplitude is $\sigma(r) = A(r)\exp[-iF(r)]$, $A(r) > 0$ is the real amplitude, and $F(r)$ is the phase of the wave. The functions $A(r)$ and $F(r)$ are presented in Figs. 4b, 4c (the curves depicted in Fig. 4a are $\varphi(r) = F(r)/2$ and $\varphi(r) = F(r)/(2+\pi)$. We can see that the spiral arms consist of two clearly defined parts: (i) the very open main spirals extending outward from the bar and (ii) the nearly circular quarter-turn spirals extending from the main spirals; each of the quarter-turn spirals is terminared by radially aligned oscillations made of short alternating segments of leading and trailing spirals. These two parts differ strongly in both their amplitudes and the degree of winding of their spirals.

Figure 5 shows the same characteristic spirals in the normal galaxy NGC 3631. In this case, the spirals exist not only in the outer region (as in NGC 1365) but also in the central region. Note that very similar spirals are observed in other galaxies as well, such as NGC 157.

As was demonstrated in Sections 2–4, we are actually dealing with a very general phenomenon: the quarter-turn spirals represent the response of the galactic disk to the gravitational potential of the main spiral, which displays characteristic behavior near the end of this spiral (which can be described as a transition from a spiral to a multipole potential regime). In our view, the close resemblance between the observed (Figs. 4, 5) and theoretical (Figs. 1, 6, 7) quarter-turn spiral patterns provides convincing support for this interpretation.

Recall that we are considering only nonresonance responses of the disk, taking the region of formation of the quarter-turn spirals to be fairly far from the main galactic resonances (or assuming these resonances to be saturated). When constructing the responses presented in Fig. 6 (NGC 1365), we used the rotation curve from [13] and the value of Ω_p corresponding to corotation, $r_{\rm CR} \simeq 210''$, and the outer Lindblad resonance, $r_{\mathrm{OLR}} \simeq 330''$. The values $r_{\mathrm{CR}} \simeq 43''$ and $r_{\mathrm{OLR}} \simeq 145''$ are well determined for the rotation curve we used to compute the response in NGC 3631 (taken from [10]). Even the presence of the inner Lindblad resonances in the galaxy is uncertain. This is due to the fact that the low accuracy of velocity measurements for the central regions of the galaxy allows variations of the rotation curve in these regions over a wide range. For example, rather smooth rotation curves for which the inner Lindblad resonances are completely absent (for any reasonable values of Ω_p) are allowed by the available measurements. It stands to reason that these comments on the uncertainty of the rotation curve in the central region are also relevant for the SB galaxy NGC 1365; however, this is not important, since in that case the quarter-turn spirals form only in a region that is fairly far from the center, outside the main spirals. For simplicity, we used a rotation law for which the inner Lindblad resonances are absent in our computations of the response in the inner region of NGC 3631. However, we emphasize again that the appearance of the quarter-turn spiral responses varies little, even under strong variations of the equilibrium parameters of the disk (and, in particular, of the rotation curve and the value of Ω_p). We will make some additional comments in this connection in the Conclusion.

6. CONCLUSION

(1) Thus, we have studied certain characteristic features of the behavior of the potential outside the region of the main spiral arms in detail and shown that the galactic disk responds to this potential by forming quarter-turn spiral structures. The main property of the potential's behavior is its rapid restructuring from a spiral to a multipole form. The azimuth $\varphi_>^{(2)}$ of the outer asymptote of the curve loci of the minima of the total potential is an important characteristic of the mass distribution in the galaxy; no less important information is contained in the value of the azimuth $\varphi_<^{(2)}$ of the inner asymptote. We emphasize that these azimuths can be determined directly from photometric observations: they differ by $\pi/2$ from the corresponding azimuths in the surface-density distribution (Figs. 4a, 5a), which can be derived from the brightness distribution of the stars or gas in any wavelength range (although our model of a disk with circular orbits should give more accurate results the cooler the subsystem used). The accuracies of the azimuths $\varphi_<^{(2)}$, $\varphi_>^{(2)}$ ($\phi_a^{(1)}$ in Figs. 4a, 5a) derived from the observed azimuths $\varphi_<^{(2)} \pm \pi/2$, $\varphi_<^{(2)} \pm \pi/2$ ($\phi_a^{(2)}$ in Figs. 4a, 5a) of the asymptotes curve of loci of the brightness (density) maxima are limited by the spread of the oscillations of this curve and by the fact that the rapid falloff of the potential with distance from the spirals fairly rapidly makes the responses comparable to the background noise level.

The position of the azimuth $\varphi_>^{(2)}$ for an SB galaxy derived from observations imposes appreciable constraints, for example, on the relative masses in the spirals and bar. For example, if the bar potential is strong enough to considerably change the azimuth $\varphi \simeq \varphi_>^{(2)}$ at which the total potential acquires a multipole form, the end of the response to this potential would be shifted by this same angle (in conflict with observations). Note that the position of this azimuth for NGC 1365 computed from the B brightness distribution is close to that required by observations of the quarter-turn spiral structure (Fig. 4a). For this reason, the disk response to the potential created by the stars radiating in the B band is very similar to the observed quarter-turn spirals, in spite of the fact that the B image is, to an appreciable extent, dominated by Population I stars, which constitute a small fraction of the total mass. Unfortunately, we do not have sufficiently good images of this galaxy in other wavelength bands that simultaneously show both the main spirals and the quarter-turn spirals (we were able to use an obviously more representative and higher quality R-band image for NGC 3631).

Strictly speaking, it would be more correct to compute the quarter-turn spiral responses from the distribution of the total mass, derived from images in the B, R, I, etc., bands (with corresponding weights). However, we have seen that even using the B-band distribution alone can yield a reasonable disk response. The reason for this (e.g. in the case of NGC 1365) is probably that the brightness (density) distributions in different bands differ appreciably only in the region of the bar, far from the quarter-turn spiral structure. This is partially confirmed by available VLT data for the brightness distribution of NGC 1365 in the B and R bands. As indicated above, these images cover an insufficiently large field, since they are cut off at $R \simeq 220''$: they do not include the quarter-turn spirals and the ends of the main spirals. The responses (i.e., quarter-turn spirals) computed from these data are virtually identical for the B and R bands; in this case, this is indeed due to the fact that the amplitudes of the brightness distributions in the B and R images differ significantly only in the region of the bar. These VLT-based responses are roughly 20° closer to the ends of the bar than the response presented in Fig. 6. We suggest that the latter is closer to the true response, since it was based on higher quality data.

It is believed (see, for example, [11]) that the most adequate approximation to the real mass distribution in a galaxy is given by images in the near infrared.

However, it follows from the above discussion that the use of such images on their own is insufficient for a correct determination of the disk response in the region of the quarter-turn spiral structure. For example, these images may underestimate the total mass in the main spirals, which would have much worse consequences than inaccurately determining the mass in the bar (since the main arms are much closer to the region of the quarter-turn spirals). Using fully realistic bar and spiral mass distributions cannot strongly change the total $m = 2$ potential in the region of the quarter-turn spirals from those presented in Fig. 6 for NGC 1365 and Fig. 7 for NGC 3631.

(2) We emphasize that we have studied the two-arm ($m = 2$) Fourier harmonics of the density, which do not precisely follow the real spiral arms. The real arms are not completely symmetrical and display some irregularities (note, for example, that the lower arm of NGC 3631 appears somewhat more powerful and longer than its upper arm). Thus, the real density distribution has been replaced with an averaged and smoothed distribution. The use of Fourier transforms is the simplest way to enhance the observational data, making them suitable for comparison with the results of theoretical studies. The term enhancement for data was proposed by Elmegreen [12], who used a more complex procedure to generate symmetrical images.

Of course, the simple method we have used to symmetrize the images is obviously not optimal, neither in general nor with regard to specific features. For example, dividing the galactic disk into rings and determining the brightness maxima in these rings is, generally speaking, poorly suited for distinguishing nearly circular structures (such as the quarter-turn spirals). The procedure used above in which a single brightness maximum is determined in each ring can also lead to incorrect results when the ends of the main spirals and some part of the quarter-turn spirals lie at the same radii (but at different azimuths). For these reasons, to increase the trustworthiness of our identification of the quarter-turn spiral structure in NGC 1365, we found it necessary to add a slight inclination to its disk, and we used the inclination $i' = 54°$ in place of the intrinsic value $i = 48°$ during the deprojection. In the case of nearly circular structures, it is probably more natural to look for brightness maxima not in rings but along radial sections. On the other hand, our method of determining maxima in rings is probably nearly optimal for distinguishing the radially oriented multipole ends of the quarter-turn spiral structures. It is clear that the method used to determine the locations of the quarter-turn spirals should be appreciably improved in future studies.

Note that we have restricted our consideration here to the curves of loci of the potential minima and density maxima. We have not considered the distribution of the amplitude along the quarter-turn spirals, although it can be used to derive additional information about both the mass distribution in the spirals (and bar) and the equilibrium parameters of the averaged axially symmetrical disk. It is of interest for future studies to investigate methods for extracting such information.

(3) Let us now make some comments concerning our freedom to choose one or another rotation curve and wave angular velocity Ω_p (or the corresponding corotation radius r_{CR}) when computing the disk response (see discussion above, for example, at the end of Section 5). Although this freedom is fairly broad, it must also be retained within a certain framework, and failure to do this can lead to errors. For example, it seems clear that the value $r_{\mathrm{CR}} = 145''$ proposed for NGC 1365 by Lindblad [13] falls outside the required framework. According to the rotation curve from [13] (which we have used), the outer Lindblad resonance should be $r_{\mathrm{OLR}} \simeq 210''$–$215''$. However, with this r_{OLR}, the response presented in Fig. 8a would have little in common with the observed spiral pattern (compare with Fig. 4a). This type of response occurs when ε from (35) is positive; in this case, in particular, the directions for the asymptotes of the curves $\varphi = \varphi^{\Phi}_{\min}(r)$ and $\varphi = \varphi^{\sigma}_{\max}(r)$ coincide. Exotic structures reminiscent of this type of quarter-turn spiral may be encountered in some real galaxies, such as NGC 4414 [10]. Figures 8a–8c show how the response changes as Ω_p is decreased, and, accordingly, r_{CR} and r_{OLR} are increased. The response begins to acquire properties characteristic of the observed spiral pattern (Figs. 4a or 6) only when $r_{\mathrm{CR}} > 180''$.

Some uncertainty (which can sometimes become appreciable) in computing the nonresonance responses of disks is introduced by the necessity to modify expression (35) for a linear response in the vicinity of resonances if they are present in the computation region: we then consider the resonances to be saturated. It is clear that a correct solution of this problem is possible only using a nonlinear theory (which does not yet exist). When computing the responses in disk regions adjacent to the ends of the main spirals (Figs. 6–8), we applied an artificial suppression (saturation) of the resonances, assuming that the imaginary part of the frequency γ was large at resonance radii. For example, when modeling the response in NGC 1365, we assumed

$$\gamma = \gamma(r) = \gamma_0 \left\{ 10 \exp\left[-\left(\frac{r - r_{\mathrm{ILR}}}{30} \right)^2 \right] \right. \qquad (36)$$
$$\left. + \, 10 \exp\left[-\left(\frac{r - r_{\mathrm{CR}}}{45} \right)^2 \right] \right.$$

$$+20\exp\left[-\left(\frac{r-r_{\rm OLR}}{60}\right)^2\right]+1\Bigg\},$$

where $\gamma_0 = 0.15$ ($\gamma_0 \ll \Omega_p \approx 1.5$ in the units used). Note that the response depends only weakly on the width of the exponential spread of the resonances. However, if the region of radii of interest to us (corresponding to the outer quarter-turn spirals) is close to $r_{\rm OLR}$, the validity of using the described method of resonance suppression may come into doubt. In this connection, we note that the outer Lindblad radii $r_{\rm OLR} \simeq 280''$ in Fig. 8c and $r_{\rm OLR} \simeq 325''$ in Fig. 8d fall in the region of the observed quarter-turn spirals, $r > 260''$ (even in Fig. 8a, this region is not very far from $r_{\rm OLR} \simeq 210''$). However, in all cases modeled in Figs. 8a, 8c, and 8d, the regions of the disk where the response approaches its asymptotic behavior, $r > 360''$, are fairly far from the resonance. However, it is precisely the difference in the asymptotic behavior of the responses in the model in Fig. 8a and in the galaxy itself (Figs. 4a, 6) to which we wish to draw attention: these patterns are shifted relative to each other by 90°. Note that the purely resonance response (when $\gamma \equiv \gamma_0 \neq \gamma(r)$; $\gamma_0 = 0.15$) presented in Fig. 8b for $r_{\rm CR}$ from [13] ($r_{\rm CR} = 145''$, as in Fig. 8a) also differs from the observed pattern (Figs. 4a, 6), whereas the responses in Figs. 8a and 8b are qualitatively similar.

(4) We emphasize that we have primarily been concerned here with nonresonance responses. For this reason, the rotation curve in the central region was smoothed to ensure that the inner Lindblad resonances were absent for all values of the wave angular velocity Ω_p we have considered. Purely resonance responses have been investigated earlier in [18–20]. One result of these studies was the conclusion that the responses in the region of the inner inner Lindblad resonance (in contrast to all remaining responses) should resemble segments of leading spirals. Such leading spirals were recently detected in [21] (see also [22]) in photometric measurements of NGC 157. The interpretation of this phenomenon suggested in [21] is similar to that given in [18–20]: it uses exotic properties of the disk responses in the vicinity of the inner inner Lindblad resonance to perturbations associated with the short bar of NGC 157. We would like to suggest here an alternative explanation for the unusual structure detected in [21], whose main elements are segments of leading spirals. We came across similar structures on several occasions above, when modeling the inner region of NGC 3631 (Figs. 5a, 7), for example. Thus, it is possible that we are not dealing with a resonance phenomenon in the case of NGC 157 but instead with the typical behavior of a nonresonance response at the end of a quarter-turn spiral. In this case, the presence of a bar provides additional variety to the possible behaviors displayed by the response in regions closer to the disk center. In concluding this brief discussion, we note that this problem clearly requires (and is deserving of) more detailed study.

(5) Our last comment concerns the fact that coincidences between bars and spirals are observed only in early-type galaxies, but not in intermediate- or late-type galaxies [14]. The following explanation for this is proposed in [14]. The bars in early-type galaxies extend right to the corotation resonances, while the bars in late-type galaxies end long before this resonance is reached (possibly in the vicinity of the inner Lindblad resonance). Further, it is known [15] that the short-wave trailing spiral waves in a galaxy move outward beyond the corotation resonance and inward between the corotation and inner Lindblad resonances. Therefore, the "short-wave modes" produced by the bars in early-type galaxies can occupy the entire disk outside the bar, while the corresponding modes excited by the bars in late-type galaxies are not capable of propagating outward in the regions of the disk adjacent to the bar.

An obvious weakness of this reasoning is the implication that running waves (with a group velocity in accordance with the ascribed local dispersion equation) should exist in SB galaxies. In fact, we do not observe running waves in SB galaxies: the bar and spirals (both main and quarter-turn) represent components of a single normal mode. For example, the main spirals most often consist of two full (covering about 180°) arms [16], while their weaker and shorter extensions can be interpreted in a natural way as the response of the disk to the gravitational potential of the main spirals (as we have shown above). Of course, all this is very far from a picture with short waves propagating in the disk (running outward). The observed quarter-turn spirals did not arrive from some other region (for example, from the corotation region, where one could imagine they could be excited), and instead represent the direct response of the disk to the potential of the main spirals. Similarly, it is possible to demonstrate in some cases (see, for example, [17]) that the main spiral is primarily the response of the disk to an indirect (long-distance) gravitational interaction of a bar. Generally speaking, however, the self-gravitation of the main spirals can play an important role and, in general, it is more correct to speak of the interconnected components (a bar + all spirals) of a single mode.

7. ACKNOWLEDGMENTS

The authors are grateful for the possibility of using the NASA/IPAC Extragalactic Database (NED), as well as data from the ING archive. This work was supported by the Russian Foundation for Fundamental Research (project nos. 99-02-18432 and 00-15-96528).

REFERENCES

1. C. C. Lin and F. H. Shu, Proc. Natl. Acad. Sci. USA **55**, 229 (1966).
2. C. C. Lin, in *The Spiral Structure of Our Galaxy (IAU Symposium 38)*, Ed. by W. Becker and G. Contopoulos (D. Reidel, Dordrecht, 1970), p. 377.
3. C. C. Lin, C. Yuan, and F. H. Shu, Astrophys. J. **155**, 721 (1969).
4. P. A. Patsis and D. E. Kaufmann, Astron. Astrophys. **352**, 469 (1999).
5. A. Toomre, Astrophys. J. **139**, 1217 (1964).
6. F. H. Shu, Astrophys. J. **160**, 99 (1970).
7. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge Univ. Press, Cambridge, 1952; Fizmatgiz, Moscow, 1963), Vol. 2, p. 110.
8. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Nauka, Moscow, 1971; Academic, New York, 1980), p. 1039.
9. G. S. Bisnovatyĭ-Kogan and A. B. Mikhaĭlovskiĭ, Astron. Zh. **50**, 312 (1973) [Sov. Astron. **17**, 205 (1973)].
10. A. M. Fridman, O. V. Khoruzhii, E. V. Polyachenko, *et al.*, Phys. Lett. A **264**, 85 (1999).
11. B. G. Elmegreen and D. M. Elmegreen, Astrophys. J. **288**, 438 (1985).
12. B. G. Elmegreen, Astrophys. J., Suppl. Ser. **79**, 37 (1992).
13. P. O. Lindblad, Astron. Astrophys. Rev. **9**, 221 (1999).
14. B. G. Elmegreen and D. M. Elmegreen, Astrophys. J. **342**, 677 (1989).
15. A. Toomre, Astrophys. J. **158**, 899 (1969).
16. A. Sandage, *The Hubble Atlas of Galaxies* (Carnegie Inst., Washington, 1961).
17. P. Englmaier and O. Gerhard, Mon. Not. R. Astron. Soc. **287**, 57 (1997).
18. I. I. Pasha and V. L. Polyachenko, Pis'ma Astron. Zh. **19**, 3 (1993) [Astron. Lett. **19**, 1 (1993)].
19. I. I. Pasha and V. L. Polyachenko, Mon. Not. R. Astron. Soc. **266**, 92 (1994).
20. V. L. Polyachenko, in *Physics of Gaseous and Stellar Disks of the Galaxy,* Ed. by I. R. King, Astron. Soc. Pac. Conf. Ser. **66**, 103 (1994).
21. A. M. Fridman and O. Khoruzhii, Phys. Lett. A 276, 199 (2000).
22. A. M. Fridman, O. V. Khoruzhii, V. A. Minin, *et al.*, in *New Structures in Galactic Disks: Predictions and Discoveries,* Ed. by J. G. Fines and E. M. Consini, Astron. Soc. Pac. Conf. Ser. **230**, 187 (2001).

Translated by D. Gabuzda

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 16–26. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 19–30.
Original Russian Text Copyright © 2002 by Lozinskaya, Arkhipova, Moiseev, Afanas'ev.

Observations of Stellar Objects at a Shell Boundary in the Star-Forming Complex in the Galaxy IC 1613

T. A. Lozinskaya[1], **V. P. Arkhipova**[1], **A. V. Moiseev**[2], and **V. L. Afanas'ev**[2]

[1] *Sternberg Astronomical Institute,*
Universitetskiĭ pr. 13, Moscow, 119899 Russia
[2] *Special Astrophysical Observatory,*
pos. Nizhniĭ Arkhyz, Karachai–Cherkessia, Russia

Received May 7, 2001

Abstract—The single region of ongoing star formation in the galaxy IC 1613 has been observed in order to reveal the nature of compact emission-line objects at the edges of two shells in the complex, identified earlier in Hα line images. The continuum images show these compact objects to be stars. Detailed spectroscopic observations of these stars and the surrounding nebulae were carried out with an MPFS spectrograph mounted on the 6-m telescope of the Special Astrophysical Observatory. The resulting stellar spectra were used to determine the spectral types and luminosity classes of the objects. Of star we identified the only object of this spectral type in IC 1613. The results of optical observations of the multishell complex are compared to 21-cm radio observations. The shells harboring the stars at their boundaries constitute the most active part of the star-forming region. There is evidence that shocks have played an important role in the formation of the shells. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Irregular galaxies provide unique opportunities for studies of star formation triggered by the combined effect of stellar winds and supernova explosions in rich stellar groupings. Due to the absence of spiral density waves in these galaxies, the formation of giant multishell complexes around stellar groupings can proceed unhindered, enabling completion of the full ecological cycle of interaction between the stellar and gaseous components of giant molecular clouds. For the same reason, irregular galaxies display the longest scale lengths and time scales on which supernovae and stellar winds play a dominant role in the formation of new-generation stars.

IC 1613, a faint dwarf galaxy of the Local Group, provides one of the most striking examples of a giant multishell complex around a group of young stellar associations. The northeastern sector of IC 1613 harbors a prominent giant complex of ionized shells and supershells surrounding several dozen young stellar associations and star clusters. This complex is the sole site of contemporary star formation in the galaxy (see [1, 2] and references therein). The complex also includes the only known supernova remnant in the galaxy (see [3] and references therein). Comparison of the optical and 21-cm brightness distributions showed that the shells of HII are surrounded by extended shells of HI [4]. The shells of ionized and neutral gas are close to each other and partially overlap in the plane of the sky. If the sizes of the shells along the line of sight and in the plane of the sky are comparable, this indicates that these shells and supershells are in physical contact with each other. It is currently thought that collisions of massive expanding shells with each other and/or with giant molecular clouds can serve as triggers of gravitational instability and fragmentation in the collision region, leading to the formation of new-generation stars.

Narrow-band Hα images of the region considered, taken in 1995 with the 4-m telescope of the Kitt Peak National Observatory (KPNO), revealed chains of bright, compact emission-line objects located exactly at the edge of the two shells in the complex. The same region hosts association no. 17 from the list of Hodges [5] (the eastern part of this feature was later catalogued as association no. 25 in the list [2]). However, in our analysis in 1995, we were not able to establish a one-to-one relationship between compact objects and stars. Moreover, the fact that the stellar objects are located exactly along a thin shell-like structure is of considerable interest on its own. This work was motivated primarily by the desire to study the emission spectra and nature of the chain of objects and their possible relationship to triggered star formation.

With this aim in view, we performed photometry and spectroscopy of this region. Section 2 describes the instruments, observations, and data reduction

1063-7729/02/4601-0016$22.00 © 2002 MAIK "Nauka/Interperiodica"

technique employed. Section 3 describes the overall structure of the multishell complex derived from optical and (21-cm line) radio observations and identifies regions that are of greatest interest for detailed spectroscopy. Based on our photometric observations, we show that the compact emission objects in question are stars. Section 4 reports the results of spectroscopic observations of three selected fields made with the MPFS spectrograph mounted on the 6-m telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences. We obtained the spectra of individual stars located at the shell boundaries and estimated their spectral types and luminosity classes. We have determined the distribution of the intensity ratios of the principal lines in the spectra of the surrounding gaseous nebulas and constructed the gas radial-velocity field based on Hβ and [SII] line measurements. Section 5 discusses our main results and conclusions.

2. OBSERVATIONS AND DATA REDUCTION

Photometry

The shell complex was imaged in two filters on October 4–5, 2000 with the 1-m Zeiss-1000 telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences in the process of testing the new SCORPIO focal reducer. A description of the reducer and transmission curves of the interference filters used can be found at `http://www.sao.ru/~moisav/scorpio/scorpio.html`. The reducer is mounted at the Cassegrain focus of the telescope ($F/13$), and the total focal ratio of the system is $F/9$. The spectrograph uses a TK1024 1024 × 1024 CCD as a detector. The system has an angular resolution of $0.52''$/pixel and a field of view of $8.'9$. We obtained images in two intermediate-band interference filters. The filter with a central wavelength of $\lambda_c = 6620$ Å and a passband halfwidth of $\Delta\lambda = 190$ Å coadded the emission in the Hα and [NII] lines and in the continuum.

To obtain continuum images, we used a filter centered on $\lambda_c = 6060$ Å with a passband halfwidth of $\Delta\lambda = 170$ Å. The total exposure in each filter and the seeing during the observations were 1800 s and $1.5''$, respectively.

After performing standard procedures for CCD frame reduction (offset frame subtraction, flat-field correction, cosmic-ray removal), we reduced the images to an absolute energy scale using images of the spectrophotometric standard star BD+25°4655 taken on the same night. The reduction to equatorial coordinates was based on field stars whose positions were adopted from the digital version of the Palomar Sky Survey (`http://stdatu.stsci.edu/dss/dssform.html`).

Table 1. Log of MPFS spectroscopic observations

Field	Center coordinates		$T_{\rm exp}$, s	z, deg
	R.A. (2000.0)	DEC (2000.0)		
Field I	$01^{\rm h}05^{\rm m}6^{\rm s}.1$	$+02°09'34''$	2700	42
Field II	$01^{\rm h}05^{\rm m}5^{\rm s}.2$	$+02°09'47''$	2700	43
Field III	$01^{\rm h}05^{\rm m}1^{\rm s}.8$	$+02°09'35''$	2700	48

Spectroscopy

The spectroscopic observations were made on October 23–24, 2000, using the Multipupil Panoramic Field Spectroscope (MPFS) mounted at the primary focus of the 6-m telescope of Special Astrophysical Observatory. A description of the spectroscope can be found at `http://www.sao.ru/~gafan/devices/mpfs/mpfs_main.htm`.

The new spectrograph has a larger field of view, wider spectral range, and higher quantum efficiency than the earlier version of the MPFS [6]. The spectrograph uses a TK1024 1024 × 1024 CCD as a detector and enables the spectra of 240 spatial elements (in the form of square lenses) to be taken simultaneously, forming a 16 × 15 array on the sky. The angular size of an image element was $1''$. A spectrum of the night-sky background $4.'5$ from the center of the field of view is taken simultaneously. We obtained spectra with a resolution of 8 Å in the range 4350–6850 Å. The seeing was about $2''$. We observed three areas in the region of the giant shells, whose positions are shown in Fig. 2. Table 1 gives the equatorial coordinates of the field centers, the total exposure times $T_{\rm exp}$, and the mean zenith angles z at the time of observation.

The spectroscopic observations were reduced using IDL-based software developed at the SAO Laboratory for Specroscopy and Photometry of Extragalactic Objects. The preliminary data reduction included offset-frame subtraction, flat-field correction, cosmic-ray removal, extraction of the individual spectra from the CCD images, and wavelength calibration using the spectrum of an He–Ne–Ar calibration lamp.

We then subtracted the night-sky spectrum from the linearized spectra and converted the observed fluxes to an absolute energy scale using observations of the spectrophotometric standard star Feige 110 (taken immediately before observing the program objects at a zenith distance of $z = 50°$). We adopted the parameters for this standard from the public database `http://www.eso.org/observing/standards/spectra`. Our air-mass corrections were based on the mean spectral atmospheric-extinction curve for the Special Astrophysical Observatory given in [7].

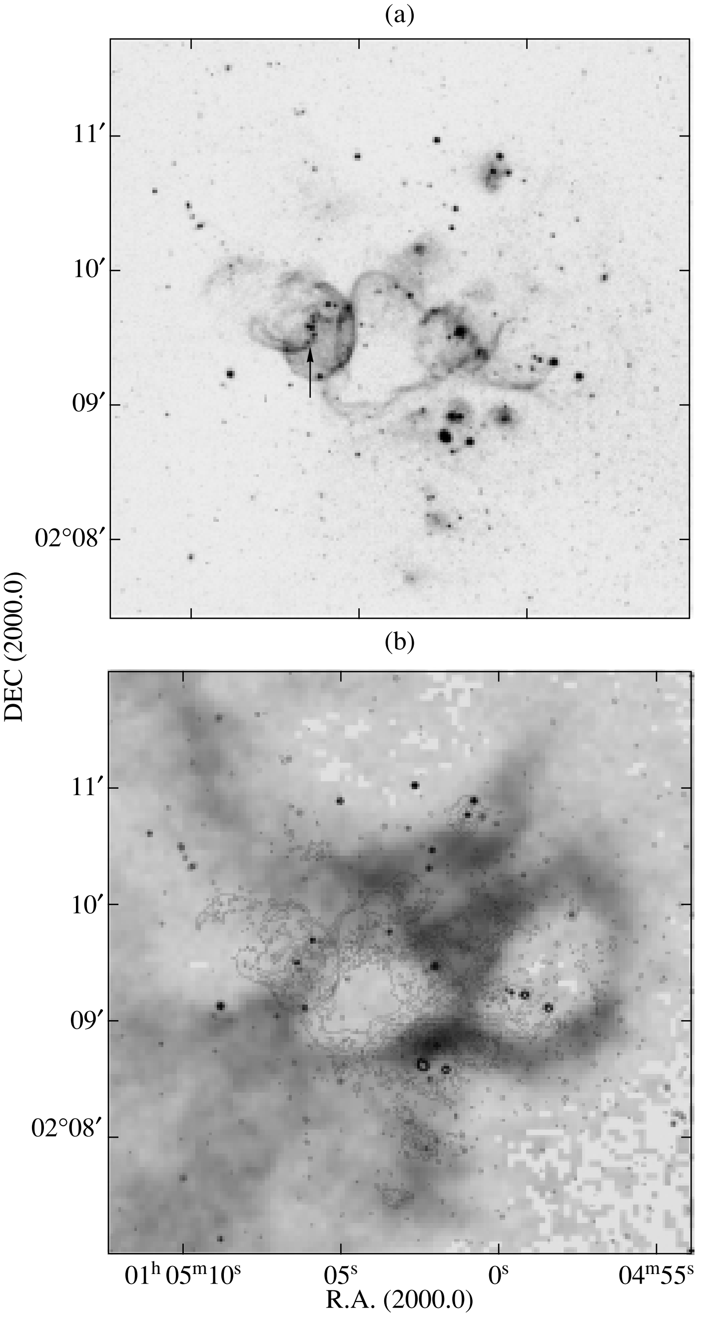

Fig. 1. (a) Hα image of the multishell complex taken with the 4-m telescope of the Kitt Peak National Observatory. The arrow shows the chain of bright, compact emission objects at the shell boundary. (b) HI brightness distribution (shades of gray) superimposed on a narrow-band Hα image of the same area (isophotes) for the northeastern sector of the galaxy IC 1613 (fragment of a chart published in [4]).

In the course of our work, we compared the Hα and [NII] line fluxes measured using the MPFS data with those derived from images of star-free shell regions taken with filters. The two methods (spectrophotometric and photometric) yielded line fluxes that agreed to within about 20%, providing an independent estimate of the actual accuracy of our observations.

3. OVERALL STRUCTURE OF THE MULTISHELL COMPLEX IN THE STAR-FORMING REGION

The multishell complex in the northeastern sector of IC 1613 is the most prominent structure seen in narrow-band Hα images of the galaxy [8–12]. This is the place where the overwhelming majority of the galaxy's bright HII regions, shells, and supershells

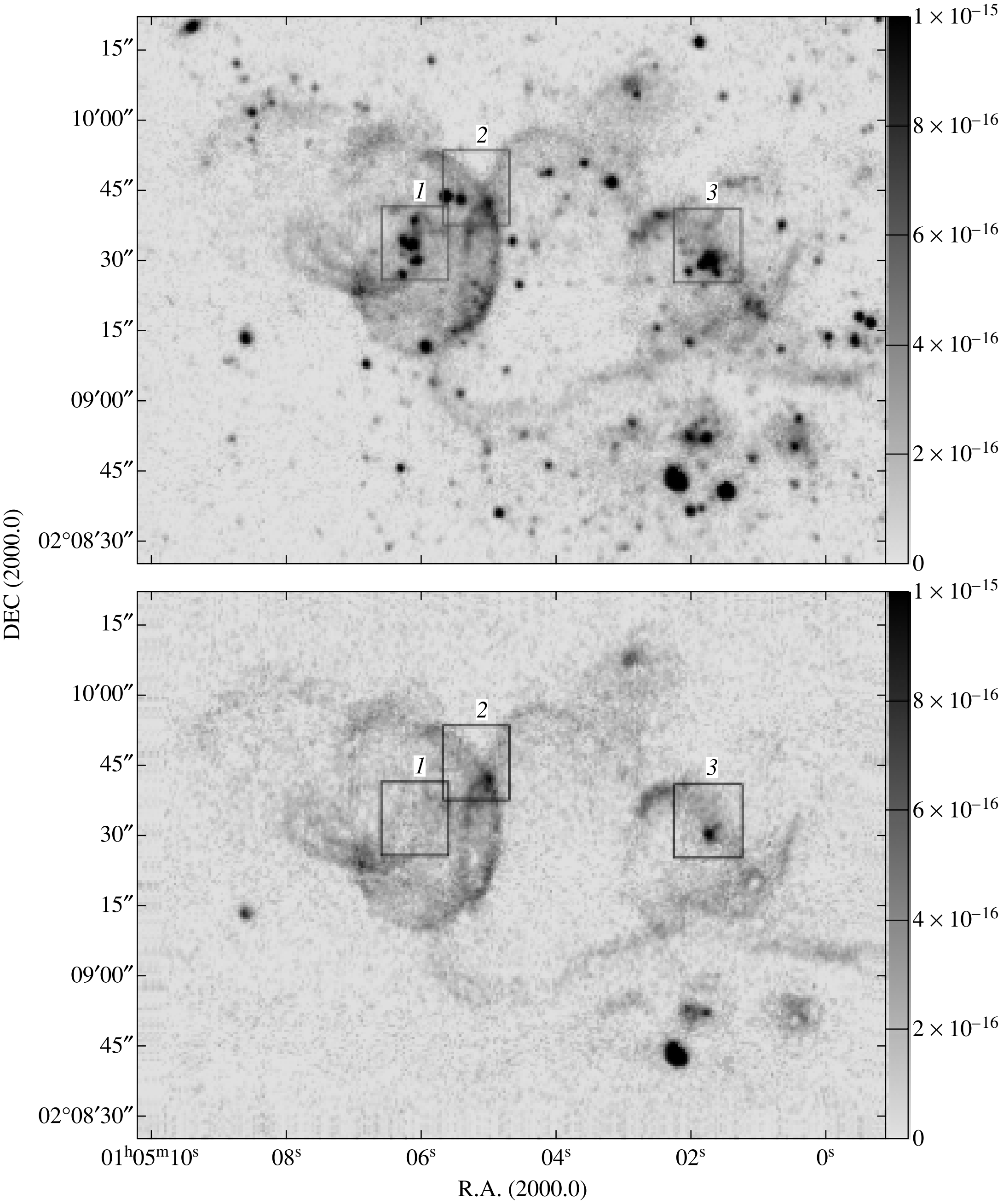

Fig. 2. Image of the complex of ionized shells taken with the Zeiss-1000 telescope equipped with the SCORPIO focal reducer. The scale is in erg/(s cm^3 arcsec). Top: image taken with a filter centered on 6620 Å and a passband halfwidth of 190 Å coadding emission in the Hα and [NII] lines and the continuum. Bottom: emission-line image (continuum subtracted). The rectangles indicate regions observed with the SAO 6-m telescope using the MPFS.

are concentrated. The stellar component of the complex is represented by about twenty stellar associations and star clusters [2, 5].

Figure 1a shows an Hα-line image of the multishell complex taken in 1995 with the KPNO 4-m telescope. The chain of bright, compact emission-line objects located exactly at the bright rim of the shell (indicated by an arrow) and several compact objects at the edges of other shells are clearly visible.

The two shells with compact objects at their boundaries correspond to objects R1 and R2 in the list of shells [12]. Both shells are located within square N27 in Fig. 3 of [12] and include the nebulae S10 and S13 in the classification of Sandage [13].

Meaburn *et al.* [1] were the first to analyze the kinematics of the complex of ionized shells. Their five spectrograms densely covered the bright part of the complex, enabling determination of the characteristic shell expansion velocities, which proved to be about 30 km/s. Valdez-Gutierrez *et al.* [12] constructed the line-of-sight velocity field of the entire complex in the Hα and [SII] lines and estimated the expansion velocities of its constituent shells and supershells. HI observations of IC 1613 [14] showed that the complex is located in the region of the brightest "spot" in the 21-cm radio emission.

Lozinskaya *et al.* [4] were the first to make a detailed comparison of the Hα and 21-cm radio brightness distributions of the complex. Figure 1b shows a fragment of the map of the northeastern sector of the galaxy published in [4], with the HI brightness distribution superimposed on the Hα line image shown in Fig. 1a.

An HI map with a high angular resolution of $7.''4 \times 7.''0$ (corresponding to a linear resolution of $\sim$25 pc) was constructed using 21-cm VLA observations (part of a large project to analyze the structure and kinematics of the neutral gas in IC 1613 [15]).

We identified the chain of compact emission-line objects and a number of other compact objects at the edges of the shells shown in Fig. 1a in our Hα line image. These compact emission-line features could be either stars located inside compact HII regions or dense clumps of gaseous shell material at the initial stage of gravitational fragmentaion and/or shock-induced compression. To elucidate the nature of these compact objects, we performed photometry of the multishell complex using the Zeiss-1000 telescope equipped with the SCORPIO focal reducer.

The upper part of Fig. 2 shows the image of the complex obtained through a filter centered at 6620 Å with a halfwidth of 190 Å. This image coadds the emission in the Hα and [NII] lines and the continuum. The lower part of Fig. 2 shows the same image after subtraction of the stellar continuum. To determine the continuum level, we used a filter with maximum sensitivity at 6060 Å and a bandpass halfwidth of 170 Å; the shape of the filter transmission curve was rectangular rather than Gaussian. We chose the coefficient for continuum subtraction to ensure the best subtraction of foreground stars. Since the central wavelengths of the two filters differ by almost 600 Å, the fact that the slope of the continuum differs from star to star becomes important. This is why many young stars inside the shell complex appear oversubtracted, due to the appreciable slopes of their continua (see Section 4). The corresponding locations are masked in Fig. 2.

A comparison of Figs. 1 and 2 suggests that the compact objects can be neither purely gaseous clumps nor stars located inside the parent cocoon.

The squares in Fig. 2 indicate Fields I, II, and III for which the spectroscopic observations reported in the next section were made. Figures 3, 4, and 5 show enlargened images of Fields I, II, and III, respectively, obtained with the same SCORPIO focal reducer with a filter centered on 6060 Å and a passband halfwidth of 170 Å. These continuum images of the three fields indicate that the compact, emission-line objects in question are stars.

We can see 13 stars in Field I, two of which are at the edge of the field. The stars in Field I belong to association no. 25 in the list of [2] (the eastern part of association no. 17 from [5]). The brightest of these stars are indeed located along the bright rim of the shell. We can also see two star-like objects at the boundary of the ionized shell in Field II. Part of association no. 17 from [5] is located in this same place. Field III contains stars from association no. 13 from the list of Hodge (1978) (no. 18 in [2]).

Below, we report the results of spectroscopic observations of the brightest of these stars made with a spectrograph mounted on the 6-m telescope of the Special Astrophysical Observatory.

4. RESULTS OF SPECTROSCOPIC OBSERVATIONS

Analysis of Field Spectroscopy

We fitted Gaussians to emission-line profiles to construct a series of monochromatic images of the fields in the Hα and Hβ lines and in the [OIII] 5007 Å and [SII] 6717, 6731 Å forbidden lines. All the line profiles can be adequately fitted by a single Gaussian, without any systematic deviations. Note, however, that our spectral resolution (350–450 km/s) substantially exceeds the expected velocities of relative gas motions [12]. In addition, we also constructed images of the program fields in the stellar continuum at wavelengths $\Delta\lambda = 4600$–4800 Å. Comparison of these data with the continuum images obtained with the Zeiss-1000 telescope enabled us to coadd the spectra from the spatial elements corresponding to individual stars.

Figures 3–5 show the results obtained for Fields I, II, and III, respectively (continuum and emission-line field images and the spectra of individual stars with their surrounding nebulae).

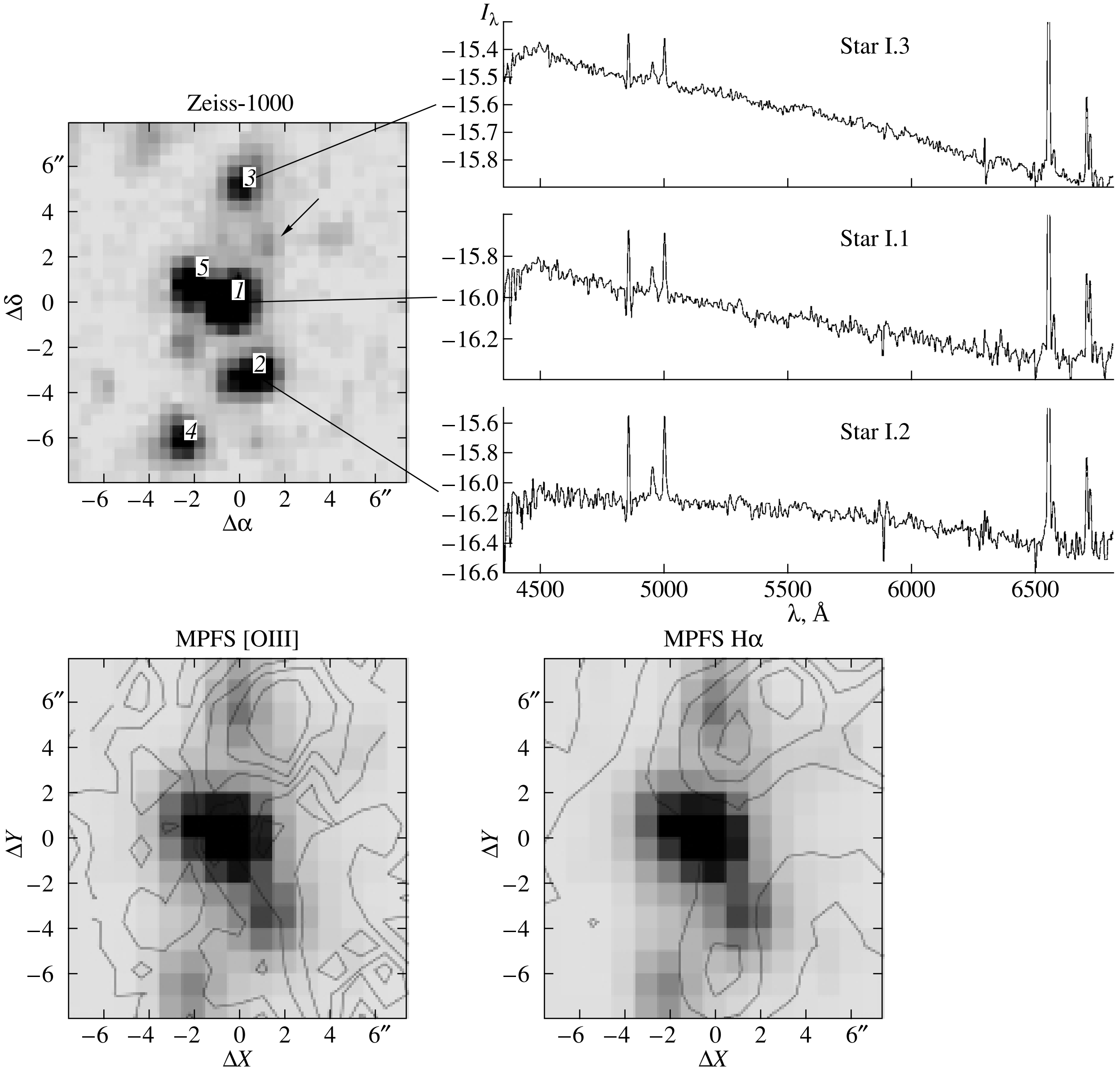

Fig. 3. Top left: Continuum image of Field 1 taken with the Zeiss-1000 telescope through a filter centered on $\lambda_c = 6060$ Å with a passband halfwidth of $\Delta\lambda = 170$ Å. The star numbers are indicated; the arrow marks the hot star that cannot be seen in the MPFS data (see text). Top right: Spectra of the three brightest stars in the field. Bottom: Isophotes of [OIII] (left) and Hα (right) line images superimposed on the 4600–4800 Å continuum images (MPFS data).

Field I

We analyzed the spectra of five stars in Field I to estimate their spectral types. Unfortunately, the low resolution and fairly low signal-to-noise ratio prevented us from reliably detecting all spectral features. In addition, superimposed emission lines from ionized gas in the vicinity of the stars can be seen in all the spectra.

Star I.1 can be classified as an O supergiant, judging from the HeII 4541, 4686 Å and SiIII 4552 Å lines in its spectrum.

Star I.2 is evidently much hotter than star I.1, since it exhibits well defined SiIV 4631, 4654 Å lines; the intensity of its HeII 4686 Å line indicates that the star is a giant rather than a supergiant.

No clearly detectable absorption lines could be found in the spectra of stars I.3 and I.4.

The spectrum of star I.5 exhibits a strong HeII 4686 Å line, indicating the it is an O star of luminosity class III–V.

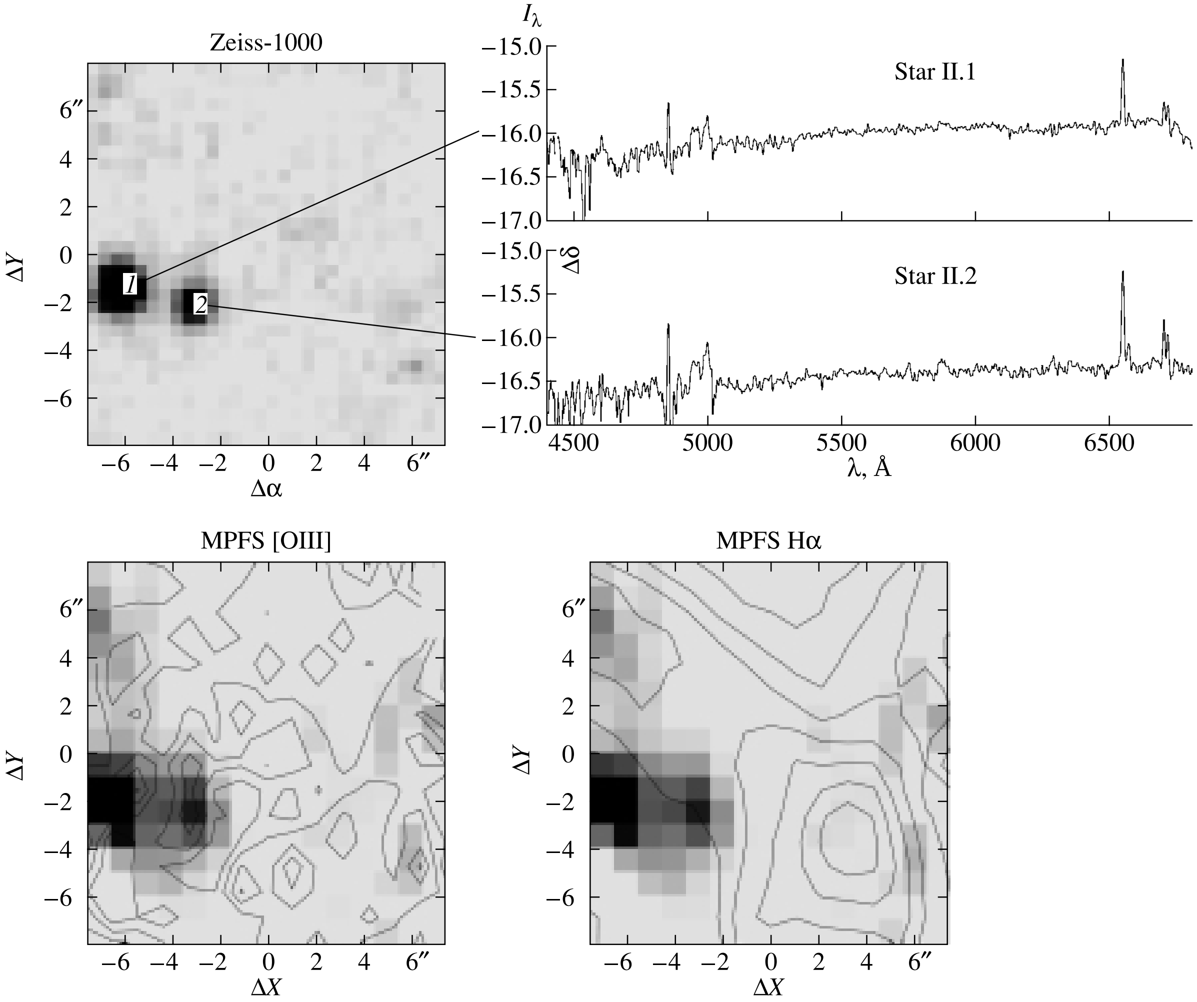

Fig. 4. Same as Fig. 3 for Field II.

We estimated the B and V magnitudes of the stars in the UBV system from their absolute spectral energy distributions, based on the calibration of [16]. The results are summarized in columns 1–3 of Table 2. Column 4 gives the $E(B{-}V)$ color excesses of the stars derived from the Hα/Hβ Balmer decrements of the nearest HII regions located within Field I. Columns 5–7 give the $(B-V)_0$ color index, absolute magnitude M_v, and the spectral type estimated from $(B-V)_0$, taking into account the luminosity classes estimated directly from the spectra.

All estimates in this paper are based on the adopted distance modulus of $(M-m)_0 = 24.31$, corresponding to a distance of 730 kpc, in accordance with the new determination [17].

The hottest star in Field I appears to be star I.5.

The emission-line spectra of the gaseous environments of each star are typical of HII regions. The emission-line ratios in the neighborhoods of stars I.1, I.2, I.3, and I.4 are I([OIII]5007/Hβ) $= 0.8$; the relative intensities of the same lines for star I.5 yield a ratio of 1.05, confirming the higher temperature of this star.

Note that all five stars in Table 2 show appreciable interstellar extinction.

There is another hot star with coordinates $\Delta X = -5''$, $\Delta Y = +2''$ between stars I.1 and I.3. This star does not show up against the bright nebular background emission in images of Field I based on MPFS data. The [OIII] 5007 Å and Hα brightness maps (see Fig. 3) show a well defined emission maximum at this location. The star can be seen on the KPNO plate as a compact clump of Hα emission and is clearly visible on the Zeiss-1000 plate. This star has $V = 21.3^m$; the extinction estimated from the Balmer decrement is low, and its temperature is high: the line intensities are I([OIII]5007)/I(Hα) $\simeq 0.3$ and I([SII])/I(Hα) $\simeq 0.2$. We therefore conclude that the excitation of the HII region surrounding the star is radiative, and not associated with shocks. The star is responsible

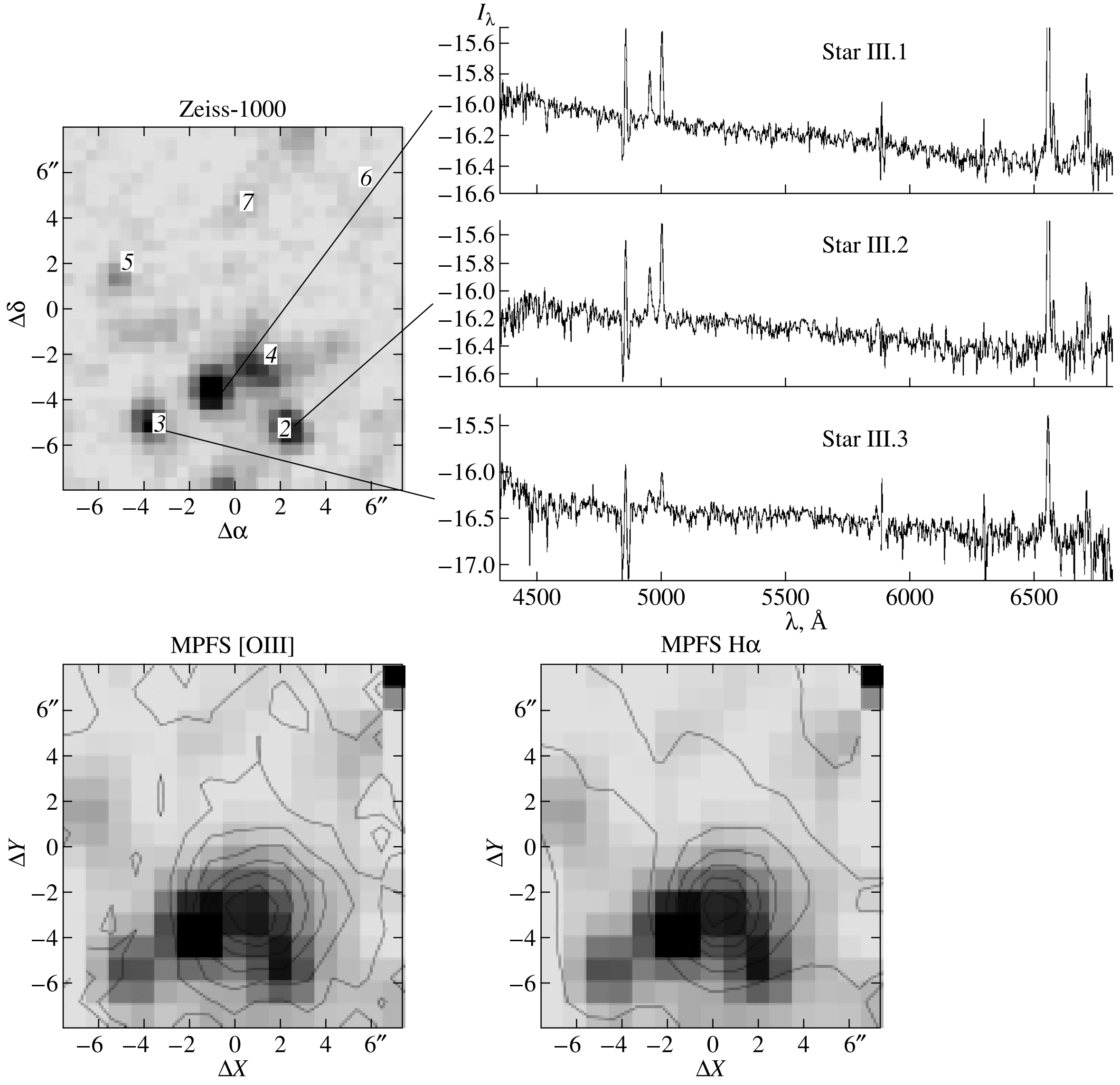

Fig. 5. Same as Fig. 3 for Field III.

for exciting both its own surrounding gas and the gas in the vicinity of the cooler star I.3.

We used the Zeiss-1000 images to estimate the V magnitudes of fainter stars in Field I relative to star I.4 (for which Table 2 gives $V = 19.8^m$). The V magnitudes of these stars are 20.9^m, 21.5^m, 21.6^m, 21.3^m, 21.5^m, and 21.8^m in order of decreasing brightness.

Field II

Our spectra of the two brightest stars in Field II show that these objects are yellow supergiants. We used the same method as for Field I to compute the parameters of these stars presented in Table 3, based on their absolute spectral energy distributions.

The well localized Hα and [SII] line emission maxima can be seen to the west of star II.2, in the region with $\Delta X = 0$, $\Delta Y = +6''$. It is here that the main source of ionizing radiation must be located. However, the emission spectrum of this region is by no means typical of an HII region ionized by an O star: the relative line intensities are $I([\mathrm{OIII}]\,5007)/I(\mathrm{H}\alpha) < 0.2$ and $I([\mathrm{SII}])/I(\mathrm{H}\alpha) \approx 0.35$. Such parameters are typical of cooling gas behind a shock front.

The electron density of the gas component in this region is low, as indicated by the [SII] intensity ratio of $I(6717)/I(6731) \simeq 1.7$, corresponding to $N_e < 100$ cm^{-3}.

Based on the 4600–4800 Å continuum intensity in this region, we estimate the upper limit of the V

Table 2. Parameters of stars in Field I

Star	V	$B{-}V$	$E(B{-}V)$	$(B{-}V)_0$	M_v	Sp
1	2	3	4	5	6	7
I.1	17.94	+0.02	0.33	−0.31	−7.4	OI
I.2	19.18	−0.10	0.18	−0.28	−5.7	OIII
I.3	19.32	+0.37	0.18	+0.19	−5.5	A7 III
I.4	19.80	+0.10	0.18	−0.08	−5.0	B8 II
I.5	19.80	−0.05	0.29	−0.30	−5.4	O8III or O4V

Table 3. Parameters of stars in Field II

Star	V	$B{-}V$	$E(B{-}V)$	$(B\text{-}V)_0$	M_v	Sp
II.1	18.94	+1.33	0.08	+1.25	−5.6	G8–K0 II
II.2	20.04	+1.53	0.18	+1.35	−4.8	K0–K2 II–I

Table 4. Parameters of stars in Field III

Star	V	$B{-}V$	$E(B{-}V)$	$(B{-}V)_0$	M_v	Sp
III.1	19.40	+0.01	(0.3)	−0.29	(−5.8)	OIII
III.2	19.62	+0.07	0.3	−0.23	−5.6	O–B
III.3	20.17	-0.06	0.3	−0.36	−5.0	Of ?
III.4	19.69	+0.14	(0.4)	−0.26	−5.8	OIII
III.5	21.2	<0?				Possibly O
III.6	21.24	Red				
III.7	21.07	Red				

magnitude of the possible ionizing star to be 19.3^m. This candidate exciting star can be seen at Δ X = 0, Δ Y = +8″ on the map obtained with the Zeiss-1000 telescope.

Note that the line-intensity ratio is $I([\mathrm{SII}])/I(\mathrm{H}\alpha) \approx 0.35$ virtually throughout the whole of Field II, as is typical of shock excitation. The same field contains part of the shell R2.

Field III

We obtained the spectra of seven stars in Field III. Table 4 summarizes the results of our photometric and spectral analysis for these stars. The color indices were determined from the Balmer decrements of HII regions surrounding the stars.

The spectrum of star III.1 appears to show the HeII 4541 Å line, whereas the HeII 4686 Å line cannot be identified. We conclude that this star is not an O supergiant but instead a star with luminosity class III. We cannot, however, rule out the possibility that this might be an early O main-sequence star.

The spectrum of star III.2 exhibits a strong [OIII] 5007 Å nebular line: $I([\mathrm{OIII}])/I(\mathrm{H}\alpha) \simeq 0.4$. This may indicate a high temperature for the exciting star ($T \geq 3 \times 10^4$ K). On the other hand, the spectrum of star III.2 may contain HeI 4713 Å line absorption, as is typical of early B stars. This star could well be binary.

The spectrum of star III.3 shows a well defined emission band at 4640–4650 Å, typical of WR and Of stars. However, the absence of other lines characteristic of WR stars leads us to conclude that III.3 should be classified as an Of star.

The spectrum of star III.4 shows well defined HeII 4541 and 4686 Å lines, and possibly SiIV 4630 and 4656 Å lines as well. We classify III.4 a an O7–O8 star of luminosity class III.

Following are the only conclusions we can draw about the three fainter stars in Table 4.

Star III.5 must be blue: its spectrum exhibits well defined traces of an He 4686 Å absorption line, and its estimated $B{-}V$ color index is indicative of a high temperature.

Stars III.6 and III.7 must be red, judging from their spectral energy distributions.

The main sources of gas excitation in Field III are stars III.1 and III.4, and probably also III.2 and III.3. The Hα line emission is concentrated primarily between stars III.1 and III.4, closer to III.4. The [OIII] line emission surrounds star III.4, suggesting that this is the hottest star in the region. According to our estimates, the size of the [OIII] emission region is no less than 6″, corresponding to a linear radius of $R = 10$ pc. The radius of the Hα emission-line region is also about 10 pc.

The [SII] line intensity ratio in this region is $I(6717)/I(6731) \simeq 1.5$, corresponding to $N_e \leq 100$ cm^{-3}.

We thus conclude that all four bright stars in Field III are O stars, without a single supergiant among them. In contrast to Field I, all stars in Field III are giants according to their luminosity classes.

Ionized Gas Velocities

Our spectroscopic observations in the Hα, [SII], and [OIII] emission lines can be used to construct the line-of-sight velocity fields of the ionized gas in Fields I, II, and III. The spectral resolution of our observations is 350–450 km/s, much lower than that of interferometric observations [12]; we therefore do not discuss the results of our velocity measurements in detail here, pointing out only our main conclusions. We found systematic deviations on the order of 20–40 km/s from the mean value of $V(\text{Hel}) \sim -246$ km/s in each of the three fields. The velocities based on the hydrogen, sulfur, and oxygen line measurements made in the same region vary within the same range. The velocity variations within the regions studied do not go beyond the interval $V(\text{Hel}) \simeq -200$ to -300 km/s. (Note that, here, we refer to the velocities of the line maxima, not those of weak features, as in Section 5).

All the velocity variations in Fields I, II, and III indicated above are in full agreement with the results of interferometric observations of the complex [12].

5. CONCLUSIONS

We obtained new observations of the only complex with ongoing star formation in the galaxy IC 1613 to elucidate the nature of the chain of compact emission-line objects we had identified earlier at the peripheries of two shells of the complex in our analysis of deep Hα images. Our new continuum images obtained with the Zeiss-1000 telescope have shown that these compact objects at the boundaries of two shells are stars.

We performed spectroscopy of selected fields in the region of the complex using the MPFS spectrograph mounted on the 6-m telescope of the Special Astrophysical Observatory. We obtained spectra of the stars forming the chains and estimated their spectral types and luminosity classes. The stars in question are at different evolutionary stages.

Four of the stars forming the chain in Field I at the boundary of the shell R1 are OB stars (with luminosity classes I, II, and III), while one is a cooler object of spectral type A7 III. The position of this last star deviates somewhat from the regular chain structure; we believe that this A7 star is unlikely to be a member of the chain and is simply projected onto the region.

The two brightest stars in Field II lie at the boundary of the shell R2. Our spectra show that these are yellow supergiants. The two stars do not belong to associations identified in the galaxy.

The stars in Field III do not form any clear chain but are coincident with the brightest clump of emission at the boundary of shell R5. The four brightest stars of the seven whose spectra we obtained are O stars. The luminosity classes of all these stars indicates that they are giants, with no supergiants among them, in contrast to what we found in Fields I and II. We assigned a spectral type of O to star III.3. This star, with coordinates R.A. (2000) = $1^{\text{h}}05^{\text{m}}2.^{\text{s}}2$, DEC (2000) = $+02°09'28.6''$, is the only star of this type identified in IC 1613 (its cordinates have been measured to within an accuracy of 1″).

One of the three fainter stars must also be of type O; the other two appear to be cooler.

We have identified sources of ionization of the ambient gas in each of the fields considered.

The shells R1 and R2, at whose boundaries the chain in Field I and two stars of Field II are located, stand out among other objects of the multishell complex in a number of ways.

(1) These two shells partially overlap in the sky, and form the brightest emission-line region in the entire complex (see Fig. 1). According to estimates made in [12], the Hα luminosities of the two shells are $\log L(\text{H}\alpha) = 37.51$ and $\log L(\text{H}\alpha) = 37.77$ erg/s, respectively. The same shells appear to be the brightest in suflur lines—see Fig. 1 in [12] and Section 4 of this paper. We estimate the line intensity ratio in Field II to be $I([\text{OIII}]\,5007)/I(\text{H}\alpha) < 0.2$, $I([\text{SII}])/I(\text{H}\alpha) \approx 0.35$, typical of cooling gas behind a shock front.

(2) The expansion velocity of shell R2 (50 km/s) is the highest in the entire complex; the expansion velocity of R1 is 30 km/s [12]. Note that Valdez-Gutierrez *et al.* [12] measured both velocities as

the mean separation between the two maxima in the double-humped integrated Hα line profiles of the corresponding shells. In fact, images of the complex in hydrogen and sulfur lines in various velocity intervals show both shells to display very complex kinematics. The fact that the brightness of both objects is high even at the extremes of the velocity interval covered (-272 and -177 km/s) means that the velocities of internal motions within the shells substantially exceed the typical expansion velocities indicated above. Two faint symmetric features in the integrated line profile of shell R2 with maxima at velocities of -50 and about -370 km/s are also immediately apparent (see Fig. 9 in [12]).

By analogy with the results that one of us (TAL) obtained by studying the kinematics of supernova remnants and shells surrounding OB associations, we suggest that these weaker high-velocity line features are closer to the shock-front velocity than brighter low-velocity features. In this case, the shock velocity in the shell should be at least 100 km/s. It is possible that we observe here the same "two-component" gas kinematics as in the Galactic shell surrounding Cyg OB1, Cyg OB3, and a number of other Galactic objects [18], namely the coexistence in a single shell of bright emission at low velocities and weak emission at higher velocities.

Such high internal velocities also reflect the substantial role played by shocks in the formation of both shells at whose boundaries the stellar chains are located.

(3) Comparison of the 21-cm and Hα-line brightness observations in Fig. 1b indicates that the part of the ionized shell R1 where the stellar chain is located falls in the thinnest neutral-gas bridge connecting the two neutral shell-like structures. The characteristic morphology of the ionized and neutral gas shells indicates that they are in physical contact: the neutral bridge is closely pressed up against the ionized shell R1 from outside in the region of the stellar chain, and the two shells in this region have the same radii of curvature. This morphology could be the result of the expanding shell R1 colliding with the neutral bridge located between the two HI shells.

All the facts listed above suggest that the shells R1 and R2 are located in the most dynamically active part of the star-forming complex. This is indeed where new-generation stars are most likely to be born. We plan further detailed high spatial- and spectral-resolution studies of the kinematics of the neutral and ionized gas components in this region.

6. ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-02-16118) and the Astronomy State Science and Technology Program (project 1.3.1.2). We are grateful to E. Blanton for providing an Hα-line image of IC 1613 taken upon our request with the 4-m telescope of the Kitt Peak National Observatory, to V. N. Komarov for his assistance with the observations at the Zeiss-1000 telescope, and to the Program Committee of the 6-m telescope for providing observing time.

REFERENCES

1. J. Meaburn, C. A. Clayton, and M. J. Whitehead, Mon. Not. R. Astron. Soc. **235**, 479 (1988).
2. L. Georgiev, J. Borissova, M. Rosado, *et al.*, Astron. Astrophys., Suppl. Ser. **134**, 21 (1999).
3. T. A. Lozinskaya, O. K. Silchenko, D. J. Helfand, *et al.*, Astron. J. **116**, 2328 (1998).
4. T. A. Lozinskaya, A. V. Moiseev, V. L. Afanas'ev, *et al.*, Astron. Zh. **78**, 485 (2001) [Astron. Rep. **45**, 417 (2001)].
5. P. W. Hodge, Astrophys. J., Suppl. Ser. **37**, 145 (1978).
6. V. L. Afanas'ev, V. V. Vlasyuk, S. N. Dodonov, *et al.*, Preprint No. 54, SAO RAN (1990).
7. T. A. Kartashova and N. M. Chumakova, Astrofiz. Issled. (Leningrad) **10**, 44 (1978).
8. J. Lequeux, N. Meysonnier, and M. Azzopardi, Astron. Astrophys., Suppl. Ser. **67**, 169 (1987).
9. P. Hodge, M. G. Lee, and M. Gurwell, Publ. Astron. Soc. Pac. **102**, 1245 (1990).
10. J. S. Price, S. F. Mason, and C. A. Gullixson, Astron. J. **100**, 420 (1990).
11. D. A. Hunter, W. N. Hawley, and J. S. Gallagher, Astron. J. **106**, 1797 (1993).
12. M. Valdez-Gutierrez, M. Rosado, L. Georgiev, *et al.*, Astron. Astrophys. **366**, 35 (2001).
13. A. R. Sandage, Astrophys. J. **166**, 13 (1971).
14. G. Lake and E. D. Skillman, Astron. J. **98**, 1274 (1989).
15. E. Wilcots *et al.* (2002) (in press).
16. V. Straizys, *Multicolor Stellar Photometry* (Mokslas, Vil'nyus, 1977; Pachart Publ. House, Tucson, 1992).
17. A. Dolphin, A. Saha, E. D. Skillman, *et al.*, Astrophys. J. **550**, 554 (2001).
18. T. A. Lozinskaya, Pis'ma Astron. Zh. **24**, 285 (1998) [Astron. Lett. **24**, 237 (1998)].

Translated by A. Dambis

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 27–33. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 31–37.
Original Russian Text Copyright © 2002 by Mashonkina, Shavrina, Khalack, Polosukhina, Tsymbal, Vygonets.

Effects of Spot Structure of Lines of Rare Earths and Non-LTE Effects on Lithium Abundance Estimates for Two roAp Stars

L. I. Mashonkina[1], **A. V. Shavrina**[2], **V. Khalack**[2], **N. S. Polosukhina**[3], **V. V. Tsymbal**[4], and **V. Vygonets**[1]

[1]*Kazan State University, ul. Kremlevskaya 18,*
Kazan, 420008 Tatarstan, Russia
[2]*Main Astronomical Observatory,*
National Academy of Sciences of Ukraine,
ul. Zabolotnogo 27, Kiev, 03680 Ukraine
[3]*Crimean Astrophysical Observatory,*
Nauchnyi, Crimea, Ukraine
[4]*Tavrian University, ul. Yaltinskaya 4,*
Simferopol, Crimea, Ukraine

Received May 26, 2001

Abstract—Taking into account blending of the lithium 6108 Å line profile by adjacent rare-earth lines together with their spotted surface structure does not appreciably affect lithium abundance estimates for the atmospheres of HD 83368 and HD 60435 but provides a better fit of the observed and stimulated line profiles. Our computed non-LTE corrections reduce the lithium abundance estimates by 0.1–0.2 dex for both stars. Given the uncertainties in the lithium abundances, it is not possible to be certain whether the lithium abundances in roAp stars, or at least in their spots, exceed the cosmic (primordial) value.

1. INTRODUCTION

An analysis of lines of rare-earth (RE) elements in the atmospheres of HD 83368 and HD 60435 using the STARSP code [1] to compute synthetic spectra for the entire stellar photospheres was presented in [2]. That paper also analyzes lithium line profiles, taking into account the spotted surface distribution of lithium on rotating stars using the ROTATE code, also written by Tsymbal. The spectra used for the analysis are described in Table 1 of [2].

Our current study is aimed at estimating the effect of blending with RE lines on the profiles of the lithium 6708 Å lines, taking into account the surface spot structure for both lithium and the RE elements. We will also estimate the non-LTE corrections to the lithium abundances of both stars. In [2], we obtained the following values for lithium abundances in spots.

For HD 83368 (a model atmosphere with $T_{\rm eff} = 7750$ K, $\log g = 4.0$, [M/H] = 0 at $i = 90°$, $v_e = 35$ km/s, photospheric lithium abundance $\log(N_{\rm Li}) = 1.8$) on the scale $\log(N_{\rm H}) = 12.0$:

$$\text{spot 1—}l_1 = 173 \pm 6°, \varphi = 0 \pm 6°,$$
$$R = 33 \pm 6°, \log(N_{\rm Li}) = 3.6 \pm 0.2,$$
$$\text{spot 2—}l_2 = 337 \pm 6°, \varphi = 0 \pm 6°,$$
$$R = 35 \pm 6°, \log(N_{\rm Li}) = 3.5 \pm 0.2.$$

For HD 60435 (a model atmosphere with $T_{\rm eff} = 8250$ K, $\log g = 4.5$, [M/H] = 0 at $i = 133°$, $v_e = 11$ km/s, photospheric lithium abundance $\log(N_{\rm Li}) = 1.8$):

$$\text{spot 1—}l_1 = 11 \pm 6°, \varphi = -15 \pm 5°,$$
$$R = 44 \pm 3°, \log(N_{\rm Li}) = 3.8 \pm 0.2,$$
$$\text{spot 2—}l_2 = 205 \pm 10°, \varphi = 15 \pm 5°,$$
$$R = 40 \pm 7°, \log(N_{\rm Li}) = 2.7 \pm 0.2,$$

where l is the spot longitude, φ is the latitude, and R is the spot radius ($R = 90°$ would correspond to the entire visible hemisphere).

According to the data of Grevesse and Sauval [3], the lithium abundance in meteorites is 3.31 ± 0.04 dex. Our lithium abundance estimates for both lithium spots of HD 83368 exceed this value only slightly (3.5 dex and 3.6 dex), while our estimate for one of the spots of HD 60435 exceeds it more appreciably (3.8 dex). We attempt here to estimate the effect of blending of the Li 6708 Å line with RE lines, as well as non-LTE corrections to the lithium abundances.

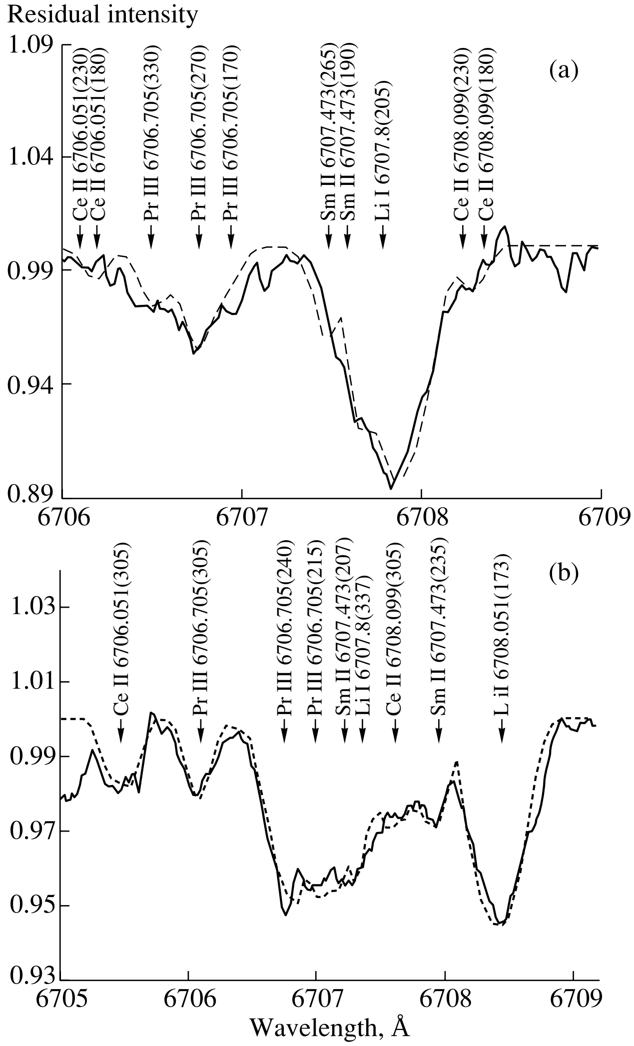

Fig. 1. Observed and computed Li I 6708 Å profiles for (a) HD 60435 (phase 0.743) and (b) HD 83368 (phase 0.689). The computations included four lines from ions of RE elements.

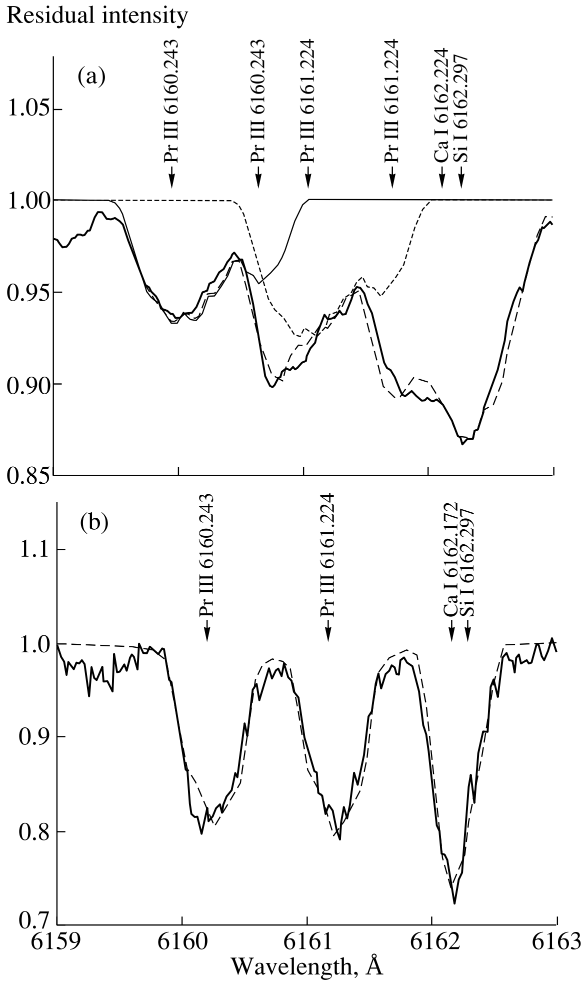

Fig. 2. Observed (thick solid curves) and computed profiles of the Pr III 6160.2 Å and 6161.2 Å lines (a) for HD 83368 (phase 0.320), taking into account two Pr III spots, the profiles of the 6160.2 Å (thin solid curve) and 6161.2 Å (short-dashed curve) lines with a large Doppler shift of the spot components due to the star's high rotation rate ($v \sin i = 34$ km/s) and the combined profile for these lines (long-dashed curve) are shown separately; (b) for HD 60435 (phase 0.005); the spot structure is not clearly seen due to the lower rotation rate, $v \sin i =$ 11 km/s.

2. MODELING THE SPOTS OF RARE-EARTH ELEMENTS

Analysis the behavior of the profiles of the lithium 6708 Å lines at different rotation phases requires allowance for blending with lines of RE elements, which also have nonuniform distributions over the stellar surface. To identify RE lines near the lithium 6708 Å line, we computed the line positions using data on RE energy levels from the NIST database. We used Kurucz model atmospheres [4] and lists of lines with atomic data from Kurucz [5] (CDROM 23) and the VALD database [6–8] to estimate the intensities of the RE lines. The atomic data for doubly ionized RE elements were taken from [9, 10]. To take blending into account, we used the intensities of the RE lines and selected those lines closest to the 6708 Å line, namely Ce II 6706.051 Å, Pr III 6706.705 Å, Sm II 6707.473 Å, and Ce II 6708.099 Å. There is no gf value for this last line, and we estimated it ($\log gf = -2.21$) from the spectrum of the narrow-line star HD 101065 (Przybylski's star, $v \sin i = 0$ km/s [13]) using the two adjacent lines Ce II 6704.524 Å and 6706.051 Å for which gf values are known. We adopted a Kurucz model atmosphere with $T_{\text{eff}} = 6750$ K, $\log g = 4.0$, and $[\text{M/H}] = 0$ for this star. Note that the Ce II lines in the spectrum of

Przybylski's star are more intense than the Li I line, in contrast to HD 83368 and HD 60435.

Figure 1 displays the spectra of HD 83368 (phase 0.689) and HD 60435 (phase 0.741), in which both lithium spots are visible, as well as the rather strong RE lines.

Our computational method using the ROTATE code is described in our earlier paper [2]. Note that this code can be used to determine the coordinates of spots based on a simultaneous solution for all rotational phases, as was done for lithium spots in [2], as well as for each rotational phase separately. The version of the code we used enabled analysis of up to five spots simultaneously, and, for the case of more than five spots, we had to solve for the coordinates of RE spots separately for each phase, considering only those spots visible at the given phase. For both stars, the accuracy of the resulting spot coordinates is no better than 8°–10°.

Nevertheless, our computations confirm, for all phases, the hypothesis that the RE elements are sometimes concentrated in ring structures around lithium spots. For example, the lithium spots of HD 83368 determined by mapping are located at the equator at longitudes of $l = 337°$ and 173°, while the rare earths are observed at $l = 20°–60°$, 100°–155°, 195°–240°, and 275°–325°. Note that we varied the latitudes for the lithium spots of HD 83368, but the best results were obtained for latitudes near zero [11]. We adopted these same latitudes for the RE elements.

Similarly, we determined the positions of RE spots on the surface of HD 60435 using the same lines of singly and doubly ionized RE elements near the Li line (Pr III, Sm II, and Ce II). An example of a computed line profile for HD 60435 at phase 0.741 taking into account the RE spots is shown in Fig. 1b. However, because $v \sin i$ (11 ± 2 km/s) is lower than for HD 83368 and due to the insufficient phase coverage, the positions of the RE spots were less accurately determined. Nevertheless, we were able to estimate the positions of the spots from the Pr III 6706.7 Å profiles for all observed rotational phases. The limiting positions of the Pr III spots (rings) from our analysis for both stars are collected in the table. Note that taking into account blending by RE lines yielded a better agreement between the computed and observed line profiles, but the lithium abundances in the spots (Section 1) remained essentially unchanged.

Our results of mapping the Li and Pr III spots of HD 83368 are in good agreement with those obtained later by Kochukhov [12], applying the Doppler imaging technique to the spectra analyzed by us, supplemented with FEROS spectra for additional rotational phases, acquired by Brazilian colleagues.

Data on Pr spots derived from the Pr III 6707.473 Å profiles for HD 83368 and HD 60435 at all available phases

Spot number	Longitude l, dex	Latitute φ, dex	Chemical abundance $\log N$, dex
		HD 83368	
1	20–60	0	3.4 ± 0.2
2	100–155	0	2.6 ± 0.2
3	195–240	0	2.7 ± 0.2
4	275–325	0	2.8 ± 0.2
		HD 60435	
1	20–50	−15	3.7 ± 0.1
2	100–150	15	3.1 ± 0.1
3	220–250	15	3.6 ± 0.1
4	300–310	−15	3.4 ± 0.3

We also derived the locations of the spots by analyzing the profiles of the Pr III 6160.243 Å and 6161.224 Å lines for both stars. Figure 2a indicates that the minimum at $\lambda = 6160.7$ Å observed for HD 83368 at phase 0.320 can be reproduced only if the spot structure is taken into account (two spots, at longitudes $l = 55°$ and $l = 130°$). The observed Pr III 6160.243 Å and 6161.224 Å line profiles at phase 0.005 for HD 60435 are best reproduced if the praseodymium spots are taken to be at longitudes $l = 50°$ and $l = 310°$ (Fig. 2b). The positions of the praseodymium spots derived using two spectral regions, 6159–6163 Å and 6706-6708 Å are in good agreement. It is obvious from the table that the wide range of Pr III abundances for both stars at various phases (3.2–3.6 dex for HD 60435 and 2.9–3.5 dex for HD 83368 [2]) can readily be explained by the variations in the Pr III abundances in the spots (ring structures).

In our computations of the Li I 6708 Å profile, we attempted to introduce lines of the isotope ^{6}Li, but the agreement between the theoretical and observed profiles became worse, even for the ratio $^6\mathrm{Li}/^7\mathrm{Li} = 0.1$ (with the exception of one of the spots for HD 60435).

3. MODEL OF THE Li I ATOM AND TECHNIQUE USED FOR NON-LTE COMPUTATIONS

Our derived lithium abundances for the spots are close to the cosmic value (HD 83368) or even exceed it (HD 60435); accordingly, it is important to consider possible non-LTE corrections for both stars. Below,

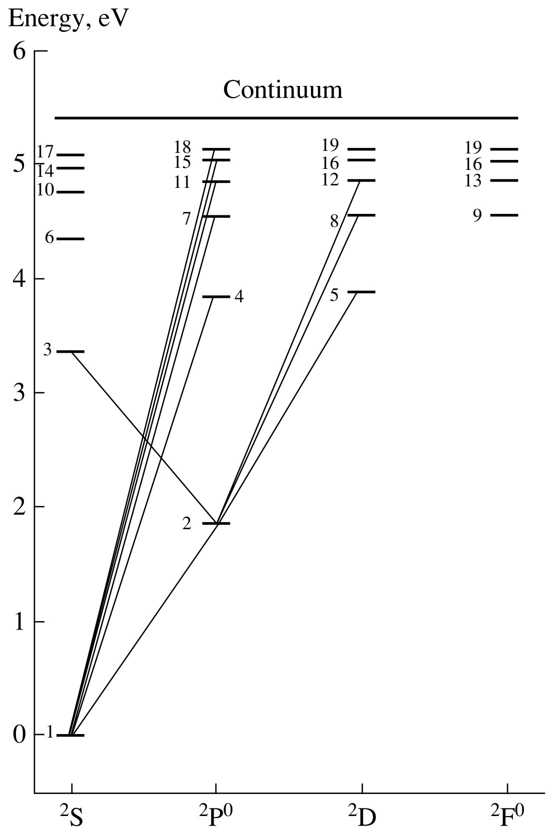

Fig. 3. Model of the Li I atom. The solid lines show linearized transitions.

we describe the computational technique we used to estimate these corrections.

If we reject the assumption of LTE, it becomes necessary to determine the populations of the atom's energy states by obtaining a simultaneous solution to the statistical-equilibrium equations for all levels of the atom, as well as to the transfer equations for all the atomic transition frequencies; we will refer to this as the non-LTE approach. The real atom, with all its possible ionization states and infinite number of excitation states, is replaced with a model atom that must include all important levels involved in transitions that can influence the populations of the upper and lower levels of the transitions under consideration. Since the stellar spectra show only one resonance Li line, 6707 Å (the $2s^2S$–$2p\ ^2P^0$ transition), the first attempts to perform a non-LTE analysis for this atom [14, 16] used a very simple atomic model, with three to four levels. Later, Steenbock and Holweger [17] considered the statistical equilibrium of Li I in cool stellar atmospheres using a ten-level atomic model. Carlsson *et al.* [18] used a 21-level atomic model to perform a more complete analysis of non-LTE mechanisms for Li I for a wide range of stellar parameters. A series of non-LTE studies of the Li I resonance line in the spectra of stars of various types have been presented by Pavlenko [19]. These investigations showed that non-LTE effects in the Li I 6707 Å line depend on the star's parameters and atmospheric lithium abundance. Until now, there have been no attempts to carry out non-LTE analyses of the formation of the Li I lines for Ap stars; this is one of our goals in the present study.

Our model for the Li I atom is the same as that used by Pavlenko [19] and includes all states for nl electronic configurations with $n \leq 7$ and $l \leq 3$, as well as the Li II ground state—a total of 20 levels (Fig. 3). In the statistical-balance analysis, we did not take into account fine level splitting and combined levels with close energy values: 6d and 6f, 7d and 7f. We took the level energies and oscillator strengths, f_{ij}, for all permitted transitions from [20]. Comparing the f_{ij} values from [20] with those computed in the Opacity Project and available in the TOP-BASE database, we find that deviations do not exceed 10% and therefore cannot significantly influence the results of the non-LTE computations. We used the photoionization cross sections, σ_{ph}, from [21] estimated using the quantum-defect method when computing the radiative rates of bound–free transitions. These coincide with the corresponding data for Li I in the TOP-BASE data base to within 10%; however, their deviation from the hydrogen-like cross sections often used for Li I [19] is appreciable. Figure 4 compares the photoionization cross sections from [21] with hydrogen-like cross sections for the six lower levels. The σ_{ph} values are close to each other near the ionization threshold, but the frequency dependences are different in all cases, with the exception of the 3p level. We made special computations to test the influence of the uncertainties in σ_{ph} on the results. We carried out non-LTE computations with hydrogen-like σ_{ph} cross sections for the set of model parameters (T_{eff}, $\log g$, [M/N], $\log \varepsilon_{\rm Li}$) = (7000 K, 4.0, 0, 3.5). The inferred deviations from LTE were nearly the same as those obtained when we used the σ_{ph} values computed using the quantum-defect method: the non-LTE correction to the lithium abundance changed by only 0.01 dex. Recall, however, that, due to the use of hydrogen-like cross sections, the uncertainties in the non-LTE abundances can be higher if we consider different stellar parameters; they reach 0.05 dex for (T_{eff}, $\log g$, [M/N], $\log \varepsilon_{\rm Li}$) = (4000 K, 3.0, 0, 3.5).

Collisional processes in the atmospheres of Ap stars are mainly due to impacts with electrons. We used equations from [22] to compute the collision velocities: the Lotz formula for ionization, the Von Regemorter formula for excitation in permitted transitions, and the Shevel'ko formula for close permitted

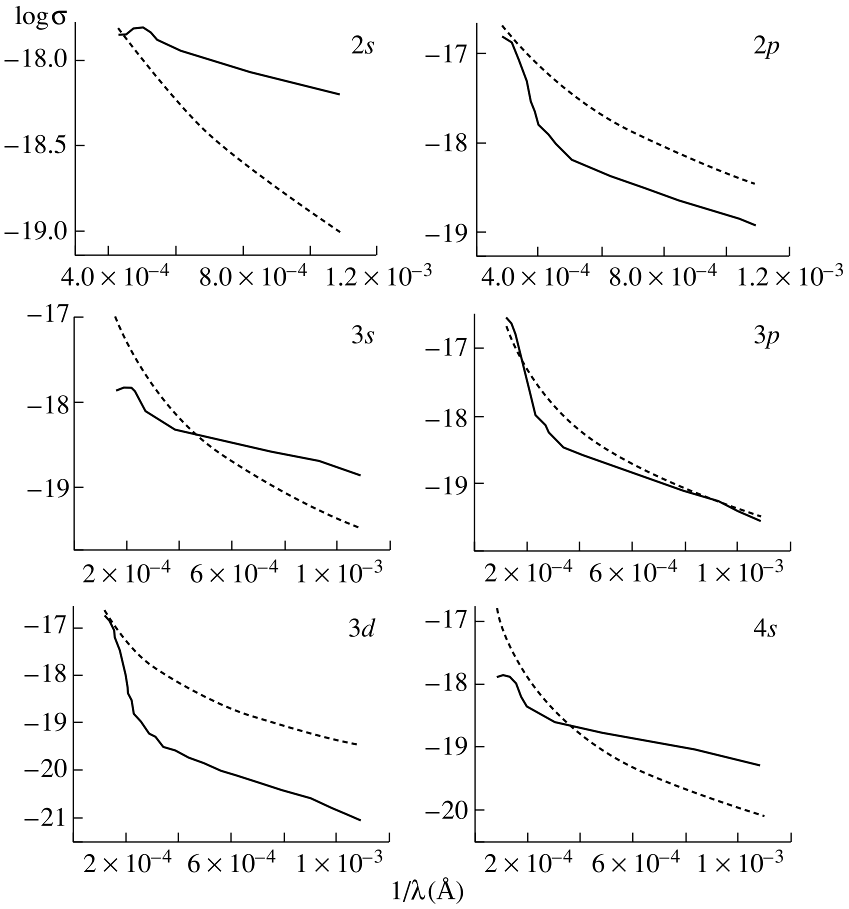

Fig. 4. Comparison of photoionization cross sections computed using the quantum-defect method [21] (solid curves) with hydrogen-like cross sections (dashed curves) for individual levels of Li I.

transitions ($\Delta E < 0.1$ eV). We took the line strength Ω_{ij} to be about equal to unity when computing the collisional excitation rates for forbidden transitions.

Our non-LTE computations used the NONLTE3 code [23], which is based on Auer and Heasley's [24] formulation of the complete linearization method. When computing opacities, in addition to the contributions from H I, H^-, H_2^+, the usual heavy elements, and electron and Rayleigh scattering, we also took into account the contributions of numerous spectral lines by summing their absorption coefficients at each frequency. Our line list was based on that of Kurucz [4] and contained about 150 000 lines with wavelengths longer than 912 Å.

In our computations of the equivalent width, W_λ, the resonance line was taken to consist of two components: 6707.761 Å with $f_{ij} = 0.502$ and 6707.912 Å with $f_{ij} = 0.251$. For natural damping, we used $\gamma_R = A_{21}$; for pressure effects, γ_6 from Gray [25] and $\gamma_4 = 10^{-8} n_e n_{eff}^5$, where n_e is the electron density and n_{eff} is the principal quantum number. To obtain model atmospheres with the required parameters, we interpolated Kurucz's grid of models [4] using code written by Pimenov, Suleimanov, and Shimanskiĭ.

3.1. Analysis of Deviations from LTE

Our computations show that the populations of the Li I energy states in the regions of the atmospheres in which the spectral lines are formed are not the equilibrium populations, leading to deviations of W_λ from those computed assuming LTE. Figure 5 displays the b factors, $\log b_i = \log n_i/n_i^*$, which describe deviations of the level populations n_i from their equilibrium values n_i^* as functions of depth in the atmosphere for various Li I levels for the (7750 K, 4.0, 0, 3.75) model atmosphere. At the depths of formation of the resonance line, in layers with $\log \tau_{1\mu\mathrm{m}}$ from 0 to −0.55, both the lower 2s and upper 2p level are overpopulated compared to the equilibrium case, with $b(2s) > b(2p)$. This overpopulation is due to the predominance of recombinations over ionizations

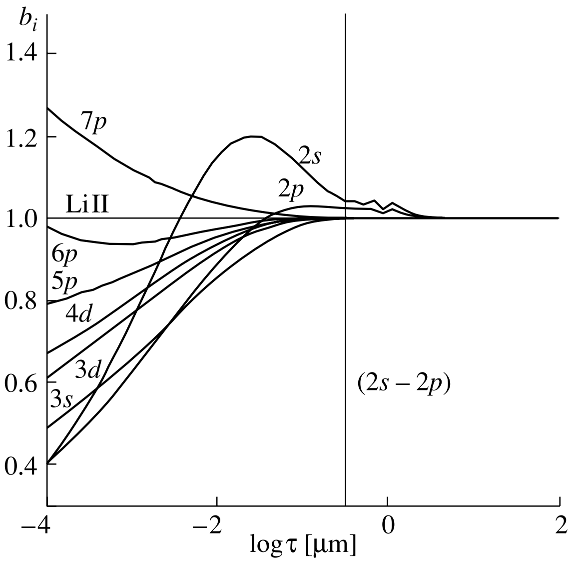

Fig. 5. Variations of b factors of the Li I levels with depth in a stellar atmosphere with (7750 K, 4.0, 0, 3.75). The vertical line corresponds to the depth of formation of radiation in the core of the resonance Li I line.

for highly excited levels with ionization thresholds in the infrared, with subsequent cascade transitions to lower levels; in the infrared, the mean intensity determining the photoionization rates, $J\nu$, is below the local value of $B_\nu(T_e)$, which affects the photorecombination rates. Photoionization from lower levels cannot compensate these processes, since the cross sections for the ground state are low (Fig. 4), whereas photoionizations from the 2p level begin to deplete it for $\log\tau_{1\mu\mathrm{m}} \leq -0.4$, when the medium becomes transparent to radiation beyond the ionization threshold ($\lambda_{thr}(2p) = 3496$ Å) and the mean intensity is above the local value of $B_\nu(T_e)$. Thus, the neutral lithium density in the region of formation of the resonance line exceeds the equilibrium value, and we can speak of "over-recombination" of Li I. This phenomenon is similar to that observed for the other alkali metals Na I [26, 27] and K I [26] and is due to the low photoionization cross sections of such atoms. Steenbock and Holweger [17] and Pavlenko [19] did not find any over-recombination of Li I, because they analyzed late-type stars, in which photoionization from the 2p level was already effective in depleting this level and the corresponding ground state in deep atmospheric layers, beginning with $\log\tau_{1\mu\mathrm{m}} \simeq 0$.

Analyses of the variations in the b factors can be used to predict the character of deviations from LTE in the line equivalent widths. In the resonance-line formation region, $b(2s) > 1$ and $b(2s) > b(2p)$. The first of these inequalities means that the line absorption coefficient is higher than in the LTE case. Due to the second inequality, the line source function, S_ν, is smaller than the corresponding Planck function, since we have $S_\nu \simeq \frac{b_j}{b_i}B_\nu(T_e)$ when $h\nu/kT_e > 1$. Both factors make the resonance line stronger compared to the LTE case. This means that the resulting lithium abundance should become lower if we reject the assumption of LTE; i.e., the non-LTE corrections to the lithium abundance will be negative for the considered range of stellar parameters.

We obtained for the non-LTE corrections to the lithium abundances −0.13 dex in the first spot ($\log(N_{\mathrm{Li}}) = 3.6$) and −0.15 dex in the second spot ($\log(N_{\mathrm{Li}}) = 3.5$) for HD 83368 ($T_{\mathrm{eff}} = 7750$ K, $\log g = 4.0$, $[\mathrm{M/H}] = 0$, $v_{\mathrm{turb}} = 2$ km/s) and −0.13 dex in the first spot ($\log(N_{\mathrm{Li}}) = 3.8$) and −0.11 dex in the second spot ($\log(N_{\mathrm{Li}}) = 2.7$) for HD 60435 ($T_{\mathrm{eff}} = 8250$ K, $\log g = 4.5$, $[\mathrm{M/H}] = 0$, $v_{\mathrm{turb}} = 2$ km/s). Consequently, our computed non-LTE corrections reduce the estimated lithium abundances for both stars by 0.1−0.2 dex. Non-LTE corrections to the barium abundances for both stars [28] and the europium abundance for HD 83368 [29] have also been determined, which also show abundance excesses. The estimated corrections for HD 83368 were 0.17 dex and 0.18 dex for barium and europium, respectively. The correction to the barium abundance for HD 60435 was 0.08 dex.

4. CONCLUSIONS

Taking into account blending of the lithium 6108 Å line by nearby RE lines, allowing for the surface spot structure of both lithium and the RE elements, did not significantly change the estimates of the lithium abundance in the atmospheres of HD 83368 and HD 60435. However, use of the ROTATE code did enable us to obtain a better fit of the observed line profiles. We attempted to determine the structure of the surface distribution of the RE elements for both stars (see table). When estimating the abundances of RE elements, it was important to take into account the spot patterns for all lines used, especially for the rapidly rotating star HD 83368: its spectrum can show overlapping of two nearby lines from different spots, forming a third and more intense line, as is the case for the Pr III 6160.2 Å and 6161.2 Å lines (Fig. 2).

We have shown that the non-LTE corrections reduce somewhat the lithium excesses of the program stars. These corrections could be important for stars with lithium abundances that are close to the cosmic value, and may eliminate the lithium excess for some roAp stars. Our computed non-LTE corrections reduce the lithium abundances of both stars by 0.1−0.2 dex. Given the uncertainties in our lithium abundances, we cannot be certain that the cosmic (primordial) lithium abundance is exceeded in the two roAp stars considered here, even in their spots. The highest lithium abundance, (3.6–3.7) ± 0.2 dex taking into account the non-LTE corrections, is obtained for one of the spots on HD 60435. Ambipolar diffusion was suggested as a possible mechanism for localized lithium enrichment in [30].

ACKNOWLEDGMENTS

We used a PC made available by the Kazan Astronomical and Geodetic Society for our non-LTE computations. L.I.M. and V.V. thank the Russian Foundation for Basic Research (project no. 99-02-17488) for partial financial support of this study.

REFERENCES

1. V. V. Tsymbal, Odessa Astron. Publ. **7**, 146 (1994).
2. A. V. Shavrina, N. S. Polosukhina, J. Zverko, *et al.*, Astron. Zh. **45**, 902 (2001) [Astron. Rep. **45**, 784 (2001)].
3. N. Grevesse and A. J. Sauval, Space Sci. Rev. **85**, 161 (1998).
4. R. L. Kurucz, CDROM13; CDROM23 (1993).
5. R. L. Kurucz, Harvard-Smithsonian Center for Astrophysics, Preprint No. 4080 (1995).
6. N. E. Piskunov, F. Kupka, T. A. Ryabchikova, *et al.*, Astron. Astrophys. Suppl. Ser. **112**, 525 (1995).
7. F. Kupka, N. E. Piskunov, T. A. Ryabchikova, *et al.*, Astron. Astrophys., Suppl. Ser. **138**, 119 (1999).
8. T. A. Ryabchikova, N. E. Piskunov, H. C. Stempels, *et al.*, Phys. Scr. **83**, 162 (1999).
9. J. Reader and Ch. H. Corliss, Line Spectra of the Elements NASA ADC, CDROM 3 (1980).
10. D. Bord, Astron. Astrophys., Suppl. Ser. **144**, 517 (2000).
11. A. Shavrina, N. S. Polosukhina, J. Zverko, *et al.*, Astron. Astrophys. **372**, 571 (2001).
12. O. Kochukhov, private communication (2001).
13. Ch. R. Cowley and G. Mathys, Astron. Astrophys. **339**, 165 (1998).
14. E. A. Muller, E. Peytremann, and R. De la Reza, Sol. Phys. **41**, 53 (1975).
15. R. E. Luck, Astrophys. J. **218**, 752 (1977).
16. R. De la Reza, C. A. O. Torres, and I. C. Busko, Mon. Not. R. Astron. Soc. **194**, 829 (1981).
17. W. Steenbock and H. Holweger, Astron. Astrophys. **130**, 319 (1984).
18. M. Carlsson, R. J. Rutten, J. H. M. J. Bruls, and N. G. Shchukina, Astron. Astrophys. **288**, 860 (1994).
19. Ya. V. Pavlenko, *The Non-LTE Formation of Lithium Lines in Late-Type Stellar Atmospheres* [in Russian], PhD Thesis (Kiev, 1996).
20. W. L. Wiese, M. W. Smith, and B. M. Miles, *Atomic Transition Probabilities* (National Bureau of Standards, Washington, DC, 1996).
21. D. Hofsaess, At. Data Nucl. Data Tables **24**, 285 (1979).
22. L. A. Vainshtein, I. I. Sobelman, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines* (Nauka, Moscow, 1979; Springer-Verlag, Berlin, 1981).
23. N. A. Sakhibullin, Tr. Kazan. Gor. Astron. Obs. **48**, 9 (1983).
24. L. H. Auer and J. Heasley, Astrophys. J. **205**, 165 (1976).
25. D. Gray, *The Observation and Analysis of Stellar Photospheres* (Wiley, New York, 1976; Mir, Moscow, 1983).
26. J. H. Bruls, R. J. Rutten, and N. Shchukina, Astron. Astrophys. **265**, 237 (1992).
27. L. I. Mashonkina, N. A. Sakhibullin, and V. V. Shimanskiĭ, Astron. Zh. **70**, 372 (1993) [Astron. Rep. **37**, 192 (1993)].
28. L. Mashonkina, T. Gehren, and I. Bikmaev, Astron. Astrophys. **343**, 519 (1999).
29. L. I. Mashonkina, Astron. Zh. **77**, 630 (2000) [Astron. Rep. **44**, 558 (2000)].
30. J. Babel, in *Peculiar Versus Normal Phenomena in A-type and Related Stars,* Ed. by M. M. Dworetsky, F. Castelli, and R. Faraggiana, Astron. Soc. Pac. Conf. Ser. **44**, 458 (1993).

Translated by N. Samus'

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 34–48. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 38–53.
Original Russian Text Copyright © 2002 by Suleĭmanova, Pugachev.

Behavior of Pulsar B0329+54 Pulse Characteristics in the Immediate Vicinity of Mode Switch Times at 111.4 MHz

S. A. Suleĭmanova and V. D. Pugachev

Pushchino Radio Astronomy Observatory, Astro Space Center,
Lebedev Physical Institute, Russian Academy of Sciences, Pushchino, Russia

Received May 11, 2001

Abstract—Times of switches from the normal to the abnormal radiation mode have been recorded in observations of individual pulses of pulsar B0329+54 using the Large Phased Array of the Pushchino Radio Astronomy Observatory at 111.4 MHz. The variations in the amplitudes of the outer components that accompany the switch to the abnormal pulse profile occurred simultaneously in only half the cases. The phase of component IV of the integrated pulse does not vary during mode switches. In half the cases, instantaneous variations of the phases of component I and the central component during mode switches may be preceeded by additional smooth variations of the phases of individual pulses occuring over several minutes. We detected a decrease in the linear polarization of the central component by, on average, 8% in the abnormal mode for the integrated pulse, due to variations in the relative intensities of two orthogonally polarized modes of the pulsar radiation. *© 2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

In spite of appreciable variations in the intensities of individual pulsar pulses, both intrinsic and due to the interstellar medium, the signal accumulated over several minutes is usually sufficient to form the stable mean (integrated) pulse profile that is characteristic of a given pulsar. However, a number of pulsars have integrated pulses that form two shapes (modes), which are both stable but have different lifetimes after their appearance. It is usual to name the mean-pulse shape with the longer lifetime (present about 85% of the time) the "normal" mode, and the other pulse shape the "abnormal" mode. The first ten "mode-switching" pulsars were discovered in 1969–1983 and, in spite of the fact that there are now more than 1000 known pulsars, no new mode-switching pulsars have been found in the past 18 years. This reflects the uniqueness of mode-switching pulsars and the importance of further investigations of these objects for our understanding of the processes occuring both at the stellar surface and in the magnetosphere.

The current paper is concerned with one of the most interesting mode-switching pulsars. The complex shapes of its integrated pulse and individual pulses, as well as its high intensity and degree of polarization, make pulsar B0329+54 an ideal object for studies of mode switching. As a consequence, each publication dedicated to this pulsar has revealed new aspects of this unusual phenomenon [1–3].

Switching between the different modes of pulsar B0329+54 was first observed by Lyne [1] at 408 MHz. The amplitudes of the outer components I and IV, which are 7 and 15% of the amplitude of the central component in the normal mode, were found to vary by factors of about two in the abnormal mode, with component I increasing and component IV decreasing in strength. It was also reported that the abnormal pulse shape was observed once during every several hours of continuous observations and was maintained for several tens of minutes. The switching times were determined to within the pulse integration time, which was roughly three minutes.

Higher frequency observations at 2695 MHz [2] showed that the abnormal pulse shape differs from that observed at 408 MHz. The new shape of the integrated abnormal pulse was given the name H mode, to distinguish it from the L mode observed at 408 MHz.

In [3], which presents the most complete summary of results for all known mode-switching pulsars, Bartel *et al.* analyze variations in the shape of the integrated pulse profile of B0329+54 over a wide range of frequencies from 410 to 14800 MHz. This analysis shows a strong frequency dependence for the shape of the abnormal-mode pulse: the variations in intensity and phase for all components of the integrated pulse in the abnormal mode have very different characters at 2.7–14.8 GHz and 410–1400 MHz. In particular, simultaneous observations at 1400 and 9000 MHz [3] indicate that, although the distance between the central and trailing outer component decreases synchronously at the two frequencies during the switch

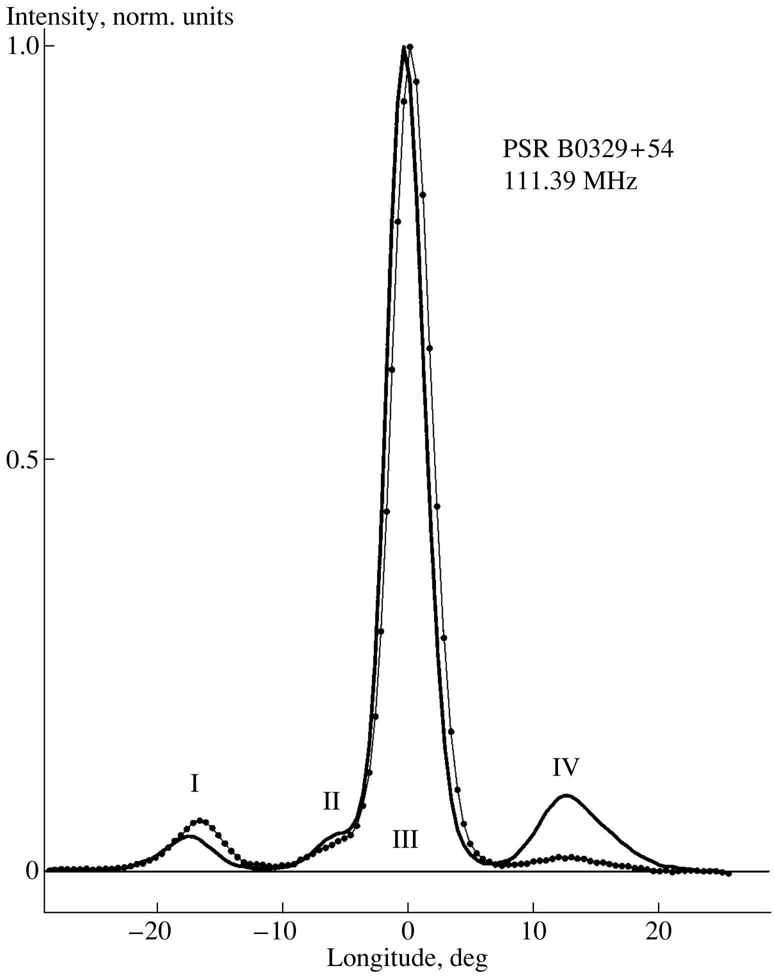

Fig. 1. Normal (thick curve) and abnormal (thin curve with dots) modes of the integrated pulse for B0329+54 at 111.4 MHz. The profiles were obtained by averaging the integrated pulses for each mode over eight days of observations and correspond to an accumulation of $8 \times 463 = 3704$ individual pulses. The pulsar period is $P = 0.7145$ s $\equiv 360°$ in longitude.

to the abnormal mode, this is due to shifts of different components at the different frequencies. Below, we will show that, at 111.4 MHz, the decrease in the distance between these components occurs purely as a result of the lag of the central component.

Our observations at 111.4 MHz broaden the frequency range for studies of mode switching in B0329+54, and are also aimed at searching for new manifestations of this effect in the properties of individual pulses. With this goal in mind, we carried out a series of measurements of varius parameters of the radiation—its total intensity, phase relative to the pulsar rotational period, and the degree and position angle of the linear polarization of individual pulses—in the immediate vicinity of mode switches. We use the term "immediate vicinity" here to reflect the fact that, with the high sensitivity of the Large Phased Array, it was not necessary to perform a preliminary integration of hundreds of individual pulsars in order to increase the signal-to-noise ratio, as was done in earlier measurements at higher frequencies [1–3].

2. OBSERVATIONS

We began a series of regular observations of pulsar B0329+54 at 111.4 MHz in June 1999 at the Pushchino Radio Astronomy Observatory in an individual-pulse recording mode with temporal resolution 0.9984 ms and a time constant of 1 ms. The observations were carried out on the Large Phased Array of the Lebedev Physical Institute using a 32-channel receiver with a 20 kHz channel bandwidth. The frequency dispersion of the pulse lag was compensated relative to the first channel at frequency 111.39 MHz. In each session of the passage of the source through the antenna beam, which lasted about 5.5 min, 463 individual pulses were recorded. Each pulse was recorded on disk and subsequently used to derive various characteristics of the signal, such

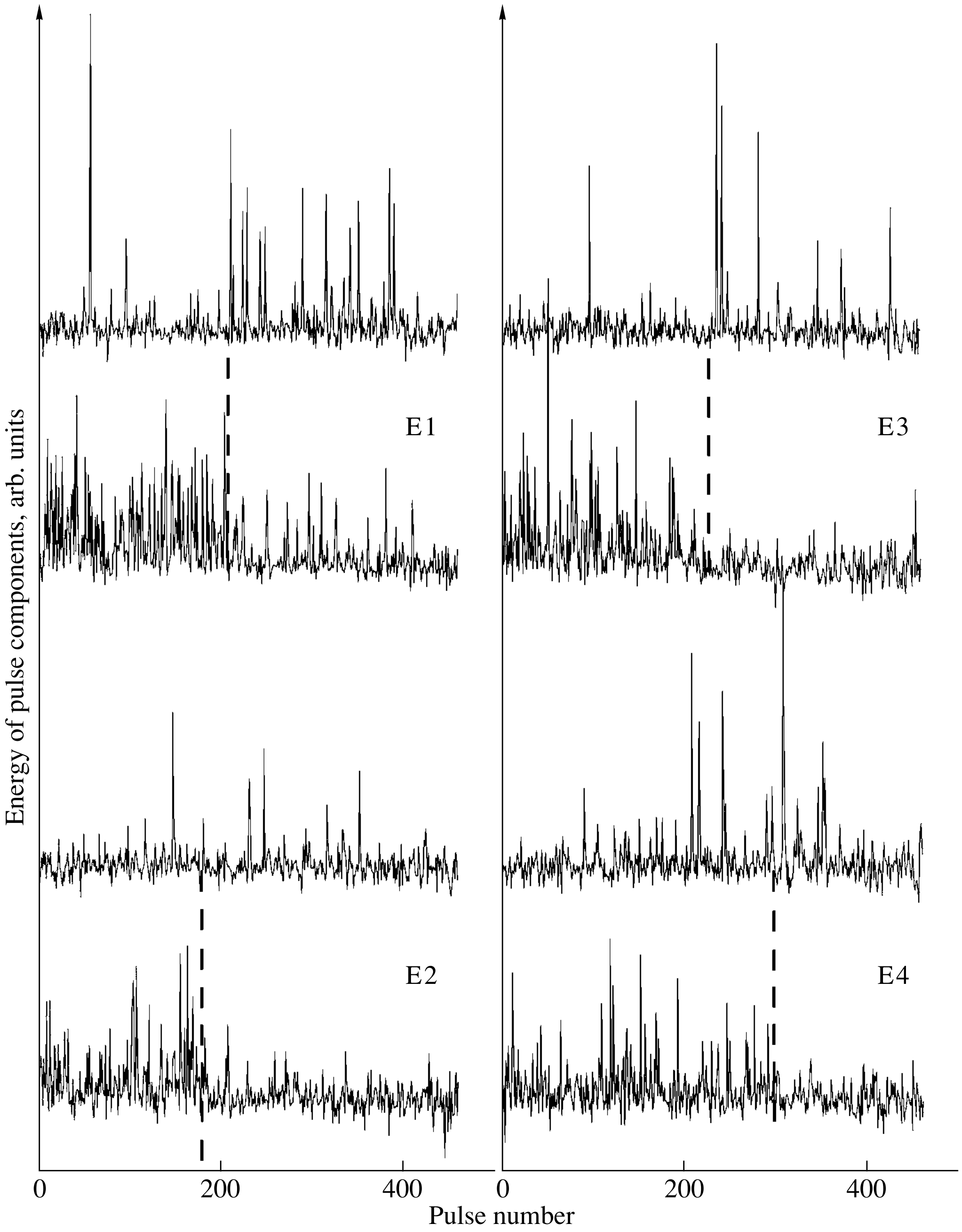

Fig. 2. Dependence of the energies of individual pulses on time (pulse number) in pairs for components I (above) and IV (below) of the integrated profiles for four mode-switching events on July 26, 1999 (E1, left), March 9, 2000 (E2, left), April 7, 2000 (E3, right), and May 19, 2000 (E4, right). The vertical dashed lines indicate the times of the mode switches.

as the phase of the subpulse in the radiation window, the total intensity, and the degree and position angle of the linear polarization. The polarization characteristics were determined from the depth and initial phase of the Faraday modulation of the amplitude at the output of the multi-channel receiver. For pulsar B0329+54, the period of the modulation at 111.4 MHz is 19×20 kHz for its rotation measure of $RM = -63.7$ rad/m^2. More detail about methods used for polarization measurements on the meter-wave antennas of the Pushchino Radio Astronomy Observatory can be found in [4, 5].

When selecting recordings for further analysis, we took special care to monitor the receiver gain. We excluded recordings in which the amplitudes of individual pulses went off scale (as a rule, at the longitude of the central component). Our analyses indicate that not performing such monitoring leads to overestimation of the amplitudes of the outer components relative to the central component and of the width of the central component. The presence of such systematic errors in [6] explains the appreciable differences in the amplitude ratios for the outer components in the integrated profile of B0329+54 at the nearby frequencies of 102.5 and 111.4 MHz. This same methodical error prevented Shabanova [6] from detecting variations in the polarization of the central component in observations of mode switches at 102.5 MHz.

3. RESULTS

Modes of the Integrated Pulse

In our series of observations from June 1999 to February 2001, which included 117 integrated profiles, we observed 24 abnormal-mode profiles (20% of the total number). Figure 1 shows the shapes obtained for the two types of pulses after additional averaging of the integrated profiles over eight days of observations, corresponding to 3700 individual pulses for each mode. The translation from time to angular measurements was done using the relation $P = 0.7145\ \mathrm{s} \equiv 360°$, where P is the pulsar period. The four main components of the integrated pulse profile can clearly be seen at 111.4 MHz. The numbering scheme for the components corresponds to that adopted by Lyne [1] in the first description of mode switching in this pulsar. Qualitatively, the intensity and phase variations of the three main components of the integrated profile during the mode switching are fully consistent with those observed at 408–1400 MHz [1, 3].

To carry out a comparative analysis of the two modes at 111.4 MHz, we separated all the recordings into two groups of pulses in the normal mode (number of recordings $N = 93$) and abnormal mode ($N = 24$). The main sign that a pulse corresponded to the abnormal mode was a significant decrease in the relative amplitude of component IV relative to the central component, compared to the situation in the normal mode. Further, for each group of pulses, we averaged the amplitudes of the outer components relative to the central component $R(I/III)$ and $R(IV/III)$ and the distances in longitude (phase) between pairs of components $S(III{-}I)$ and $S(IV{-}III)$. These measurements were done using the peak intensities for each component; the results are presented in the table.

The collected data indicate that, during the mode switching, the largest variations (a factor of 4.6) are those in the intensity of component IV. The intensity of component I varies by a factor of 1.4. The distance between component I and the central component remains constant within the errors, while the distance between component IV and the central component decreases by 1.2 ms in the abnormal mode. A comparison of the positions of the maximum of the central component in longitude on days when the mode switched within the passage time of the source indicates a systematic delay of the abnormal integrated pulse of one sampling interval equal to 0.998 ms, or 0.5°. This is in agreement with the average time delay for the abnormal-mode pulse of 1.1 ms derived from timing measurements at the nearby frequency of 102.5 MHz [6]. This is also very close to the value of 1.2 ms for the decrease in the distance between component IV and the central component detected in our observations, indicating that the position of component IV does not change during the mode switches, within the errors (±0.3 ms).

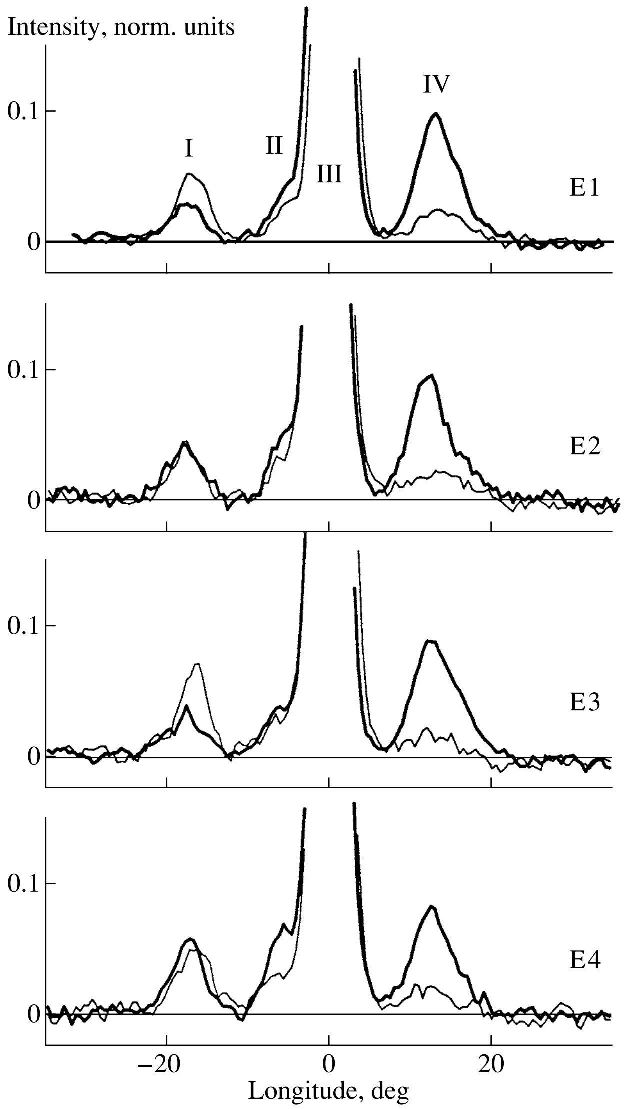

Fig. 3. Integrated pulses for B0329+54 at 111.4 MHz obtained by averaging individual pulses before and after four mode-switching events; the thick curve shows the normal mode and the thin curve the abnormal mode. We have cut off the central component at 0.15 of its total amplitude in order to better display variations in the outer components.

Characteristics of the integrated pulse of pulsar B0329+54 at 111.4 MHz

Parameter	Mean value	Standard deviation (σ)	Minimum–maximum values	Number of days
Normal mode				
S(III-I), ms	34.6 ± 0.1	1.0	32–37	93
S(IV-III), ms	24.9 ± 0.1	0.8	23–27	93
R(I/III), %	4.2 ± 0.1	0.7	2.4–6.8	93
R(IV/III), %	9.2 ± 0.1	1.3	7.1–13.1	93
DLP(0), %	40.6 ± 0.5	4.4	31–52	93
Abnormal mode				
S(III-I), ms	34.4 ± 0.1	0.7	33–36	24
S(IV-III), ms	23.7 ± 0.2	1.1	22–26	24
R(I/III), %	6.0 ± 0.2	1.1	4.3–8.1	24
R(IV/III), %	2.0 ± 0.1	0.4	1.5–2.9	24
DLP(0), %	32 ± 1.2	4.5	18–37	24

Indirect confirmation of the absence of a phase shift for component IV in the abnormal mode at frequencies below 1400 MHz can be found in [3]: it is clear in Fig. 2 of that work that the difference in the phases of component IV and the central component decreases by 2 ms during the switch to the abnormal mode at 1400 MHz. This value coincides with the phase shift of the central component toward component IV at this frequency [3, Fig. 3].

Thus, analysis of the shapes of the integrated pulse profiles for the two modes suggests that there are different phase shifts, or time lags, for different components in the profile of pulsar B0329+54.

Long-Term Variations in the Switching Rate

The probability of observing pulses in the abnormal mode during our total observing interval of 20 months was, on average, 20%. However, in shorter series of observations including approximately equal numbers of integrated profiles, the probability for appearance of the abnormal mode varied from 5 to 43%, as can be seen from the data presented below.

The probabilities for observing the abnormal mode in different time intervals were 43% in June–July 1999, 6% in August 1999–January 2000, 26% in February–June 2000, and 20% in July 2000–February 2001.

The lowest mode-switching activity in B0329+54 was observed from August 1999 through January 2000. Thanks to the enhanced mode-switching activity in July 1999, we were able to detect the time corresponding to a mode switch within the 5-minute passage of the source through the antenna beam of the Large Phased Array on July 26, 1999. We will refer to this event as E1 below. Three similar events occured in another period of enhanced mode-switching activity, on March 9, 2000 (E2), April 7, 2000 (E3), and May 19, 2000 (E4). In all cases, the mode switched in the same direction (from the normal to the abnormal mode).

Detection of the Time of Mode Switching

As a first approximation, we determined the presence of mode switching in the pulsar recordings by eye, based on the shape of the integrated profile. In recordings containing mode switches, the amplitudes of components I and IV were roughly equal and corresponded to about 5% of the amplitude of the main component.

We next analyzed the time behavior of the amplitudes of all four pulse components during the observing sessions. We used the summed amplitude for all measurements in the longitude range occupied by a given component to characterize that component's emission level; we will call this parameter the component energy. We subdivided the active window occupied by the pulsar radiation into four longitude intervals corresponding to the four main profile components. The boundaries of the intervals were based on the decrease of the component intensity to the noise level outside the pulse, or on the minima in the emission between components.

On average, the variations in the energy of component III (the central component) at 111.4 MHz do not exceed 5%. The energy of the longitude subpulse, component II, decrease in the abnormal mode, but by no more than 15%. Practically speaking, the influence of variations in the amplitudes of components II and III on the shape of the abnormal pulse profile of B0329+54 is negligibly small.

Figure 2 shows the pulse-number dependence of the pulse energies at the longitudes of components I (above) and IV (below) in pairs for the four mode-switching events at 111.4 MHz. The times of the mode switches are indicated by vertical dashed lines and correspond to pulse numbers 210 (E1), 180 (E2), 230 (E3), and 300 (E4), respectively. A comparative analysis indicates that the most similar behavior is displayed by events E1 and E3 (Fig. 2), for which the weakening of component IV and the strengthening of component I occur in accordance with the classical scenario; i.e., these variations occur virtually simultaneously, as described for the first mode-switching observations at 408 MHz [1].

The appearance of pulses obtained by averaging individual pulses before and after mode switches confirm that the observed variations in the pulse energy for B0329+54 are a consequence of its mode switching. Figure 3 shows examples of pulses for the normal (thick curve) and abnormal (thin curve) modes obtained before and after mode switches (E1, E3). Their shapes are typical of those for the normal and abnormal modes (Fig. 1). We can see a departure from the classical mode-switching scenario in events E2 and E4. In event E2, the emission level of individual pulses at the longitude of component I did not change near the time of the mode switch (Fig. 2, E2). The integrated pulses in Fig. 3 (E2) confirm this result. Moreover, there is no phase delay for component I. However, the fivefold variation in the amplitude of component IV and the shift of the central component by 1 ms suggest that there was a mode switch near pulse no. 180. There were no appreciable variations in the amplitude of component I for event E4 as well (Fig. 3, E4). A comparison of the temporal behavior for the outer components for this day [Fig. 2, E4] indicates that the increase in component I did not occur at the time of the mode switch, but instead a minute earlier, near the time of pulse no. 200. Note that, to refine the time of this mode switch, we performed four intermediate averages containing 100 pulses each. The ratios of the amplitudes of the outer components of the integrated profiles for these groups were $R(I/IV) = 0.4, 0.4, 1.3$, and 2.6. We can see that a deviation from the standard normal ratio for the integrated pulse $R(I/IV) = 0.4$ is already observed for the third group of pulses; however, the typical abnormal pulse shape with $R(I/IV) = 3.0$ and an increase in the phase of the central component by about +0.5° were observed only for pulses with numbers greater than 300.

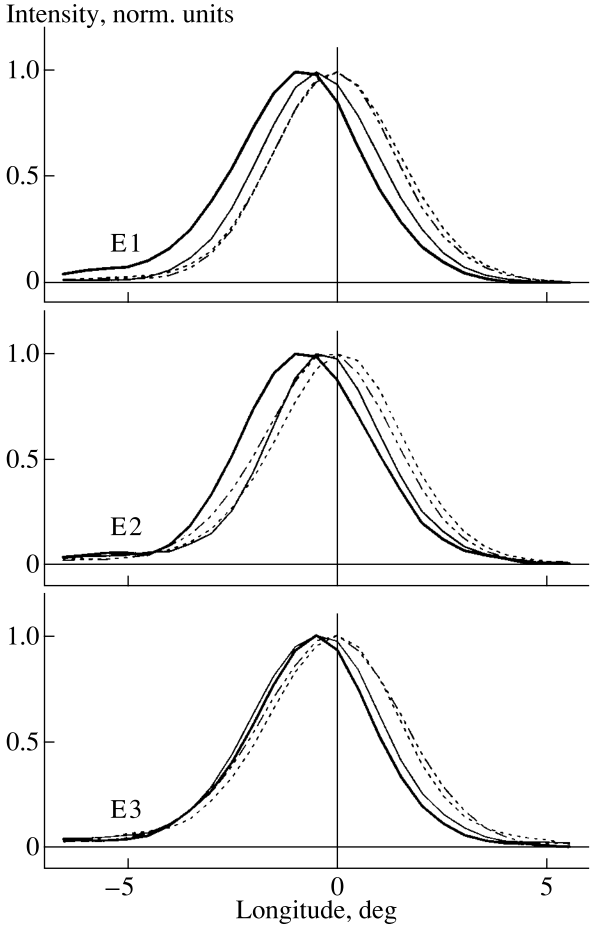

Fig. 4. Shape and position of the central component for four successive averages for events E1 (July 26, 1999, top, 105 pulses in each average), E2 (March 9, 2000, middle, 90 pulses in each average), and E3 (April 7, 2000, bottom, 115 pulses in each average). The solid and dashed curves correspond to data acquired before and after the mode switches. The position of the maximum of the averaged pulses shifts toward the right with the successive averages.

In all cases, when the intensity variations for the outer components indicate the possible presence of a mode switch, the third and deciding criterion for identifying a mode switch is the presence of a time shift of 1 ms for the central component. This shift is observed in all four mode-switching events.

It is important to emphasize that, in our case, the relative tie of pulses in intensity and phase is more accurate for observations of the two modes during a single observing session, since this excludes uncertainties associated with changes in the receiving equipment, changing conditions for interstellar scintillation, and the accuracy of relating the phases of pulses observed on different days (timing).

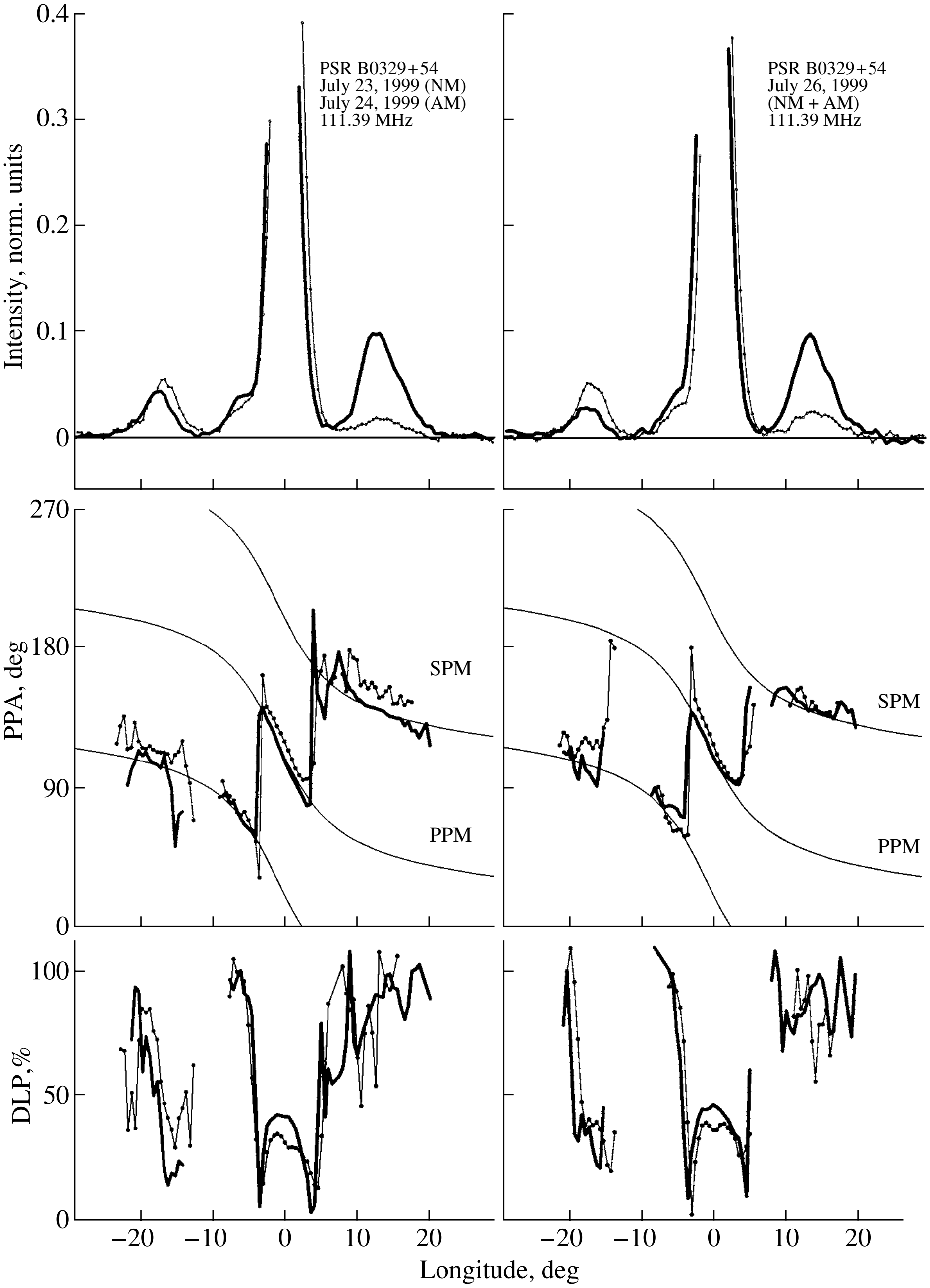

Fig. 5. Integrated profiles of the intensity, polarization position angle (PPA), and degree of linear polarization (DLP) for the observing sessions of July 23, 1999 [left, normal mode (NM), thick curve] and July 24, 1999 [left, abnormal mode (AM), thin curve], and for the two modes for event E1 of July 26, 1999 (right). For event E1, we have averaged 210 pulses before (thick curve) and 253 pulses after (thin curve) the mode switch. On the PPA profile, we have superposed a curve describing the expected behavior in a rotating-vector model with parameters $\alpha = 59°, \beta = -5°$, and $\varphi_0 = -1°$. In both cases, the degree of linear polarization at zero longitude decreased by 10% in the abnormal mode.

It follows from the above analysis that mode switches are characterized, first and foremost, by a decrease in the energy at the longitude of component IV by a factor of four to five and a time delay of approximately 1 ms for the central component. Variations in the strength of component I can occur either at the time of (events E1, E3) or preceeding (E4) a mode switch, or be absent in a time interval of several minutes surrounding a mode switch (E2).

The distinguishing feature of component I is that its emission has a flare-like character, forming an episodically appearing series of strong pulses, some of which can exceed its mean amplitude by a factor of ten. In our opinion, the growth in the amplitude of component I in the abnormal integrated pulse is not due to a systematic increase in the amplitudes of the individual pulses at this longitude, but instead to an increase in the probability of such flares, the closest of which may not coincide in time with the mode switch.

Variations in the Phases of Individual Pulses During Mode Switches

At present, long-term variations in the arrival time (phase) of pulses on time scales from tens of minutes [2, 3] to months [6] have been studied only for integrated pulses. It is usually assumed that jumps in the phases of individual pulses occur over a time that is much less than a rotational period of the pulsar [3], with constant phases preserved before and after the jump. In our investigation of the behavior of the phases of individual pulses for the four mode-switching events, we found that the phase shift by 1 ms, or 0.5°, occured in accordance with this scenario in only two events (E3 and E4). In events E1 and E2, the pulse phase did not remain constant up until the mode switch, instead smoothly growing over the course of several minutes, beginning with the first pulse in the recorded sequence and probably over some time before the start of this sequence.

This conclusion is based on the form of the time dependence of the phase of the central component. For events E1 and E2, linear fits to the data preceeding the mode switches have slopes corresponding to a rate of change of 1 deg/5.5 min. After the mode switch, the best-fit lines have zero slope. For events E3 and E4, these dependences are best fit by two lines with zero slope shifted in phase by 0.5°. We performed a similar fitting procedure for recordings that did not contain mode switches. There are no systematic variations in the subpulse phase over 463 pulses in these control recordings. Here, we should say a few words about the method we used to measure the phases of subpulses (components of individual pulses). As a first approximation, we determined the phase of the subpulses from the positions of their intensity maxima in the corresponding four longitude intervals. Obviously, the influence of random fluctuations on the accuracies of the subpulse phase centers will be largest for weak subpulses. Therefore, we selected subpulses with ampitudes exceeding a specified threshold for our analysis. Further, to decrease the influence of random intensity fluctuations on the measurements, we determined the phase of each subpulse from the position of its centroid. The subpulse phases determined in these two ways were very similar.

To obtain additional confirmation of the presence of subpulse phase drifts preceeding the mode switches, we carried out averages for four groups of pulses with 105, 90, 115, and 100 pulses for events E1 (July 26, 1999), E2 (March 9, 2000), E3 (April 7, 2000), and E4 (May 19, 2000) without selecting by amplitude. Figure 4 shows the results of this averaging for the central component for three events. The phases of the maximum of the integrated pulse in each group in milliseconds from the start of the pulsar period trace were distributed as follows: E1 (262, 263–264, 264), E2 (147, 148–149, 149), E3 (150, 150–151, 151), and E4 (153, 153, 153–154). Zero longitude in Fig. 4 corresponds to the maximum of the integrated pulse in the abnormal mode. The solid and dashed curves correspond to averages before and after the mode switch.

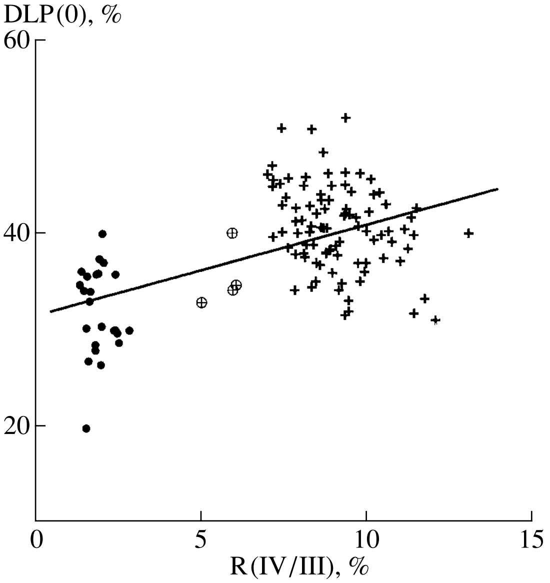

Fig. 6. Dependence of the degree of linear polarization of the central component at zero longitude on the ratio of the amplitudes of component IV and the central component for 117 observing sessions at 111.4 MHz. Data for the normal mode are represented by the 93 crosses, for the abnormal mode by the 24 dots, and for days with mixed modes by the four circled crosses.

These results demonstrate that, in events E1 and E2, there was already a shift in the pulse phase by about 1 ms between the first and second group, before the mode switch. A subsequent delay with the same magnitude is observed between pulses in the second and third group, on either side of the mode switch. The pulse phase corresponding to the abnormal mode became firmly established during the first hundred pulses after the mode switch, since the mean phase for the two last groups of pulses are the same. Thus, in half the cases, the total phase shift of the central component in the immediate vicinity of the mode switch indicated by the phase–pulse number dependence and by the pulse averaging results shown in Fig. 4 is about 2 ms, or 1°.

A growth of the phase with time preceeding the mode switches is also observed for subpulses corresponding to component I of the integrated profile. We did not make phase measurements near the times of mode switches for component IV, due to the appreciable weakening of its intensity in the abnormal mode.

Behavior of the Linear Polarization of Pulses During Mode Switches

Figure 5 presents integrated profiles of the intensity and of the position angle and degree of the linear polarization for the two modes for integrated pulses obtained on different days (left) and during a single sequence of pulses before and after the mode switch in event E1 (July 26, 1999, right). We have superimposed the curves corresponding to two orthogonally polarized radiation modes on the integrated profiles of the polarization position angle: the primary polarization mode (PPM) and secondary polarization mode (SPM). Each of these curves, which were derived from a single curve and shifted vertically by ±90°, describes the expected behavior of the polarization angles in the widely accepted rotating-vector model [7]:

$$\psi(\varphi) = \psi_0 + \arctan\left[\frac{\sin\alpha\sin(\varphi-\varphi_0)}{\cos\alpha\sin(\alpha+\beta) - \sin\alpha\cos(\alpha+\beta)\cos(\varphi-\varphi_0)}\right], \quad (1)$$

where φ is the longitude (phase), α is the angle between the rotational axis and the magnetic-dipole axis of the star, and β is the impact angle of the closest approach of the line of sight to the dipole axis. The parameters $\alpha = 59°$ and $\beta = -5°$ provide the best fit to the run of the polarization position angle of B0329+54 in longitude at 111.4 MHz and are close to the values obtained for this same pulsar based on observations at 408 MHz [8]; $\varphi_0 = -1°$ is the position of the magnetic polar axis of the pulsar on the longitude axis. Comparison of the polarization profiles for integrated pulses for the different modes shows that they are shifted in longitude together with the pulse.

Another new result of our observations of B0329+54 at 111.4 MHz is the detection of depolarization of the integrated pulse near the central longitude in the abnormal mode. We can see in Fig. 5 that, in both cases, the degree of linear polarization (DLP) at zero longitude decreased by 10% in the abnormal mode.

To verify the inferred variations in the degree of linear polarization during mode switching using a larger volume of statistical data, we carried out polarization measurements for 117 observing sessions and constructed a graph of the dependence of the degree of polarization at the central longitude on the ratio of the amplitude of component IV to that of the central component. In Fig. 6, pulses in the normal mode are shown by crosses and those in the abnormal mode by dots. The distribution of points in Fig. 6 confirms that the degree of linear polarization of the central component decreases in the abnormal mode. The correlation coefficient for the corresponding dependence is $CC = 0.5$. As expected, the DLP for four days with mixed modes (circled crosses) occupy an intermediate position between the two main groups of points in Fig. 6. On average, the variation in the DLP is $8.6 \pm 1.5\%$ ($40.6 \pm 0.5\%$ for the normal mode and $32.0 \pm 1.0\%$ for the abnormal mode).

Variation in the Relative Intensities for Orthogonal Polarization Modes During Mode Switching

The most obvious place to search for the origin of the depolarization of the integrated pulses in the abnormal mode is in variations in the relative intensities of orthogonal polarization modes. Testing this possibility requires dividing the data according to these polarization modes. We separated the initial dataset into two groups of data with polarization position angles offset by ±45° from the expected model angles for the primary and secondary polarization modes [formula (1)]. This procedure was carried out for several recordings. Figure 7 shows the longitude distribution for the polarization position angles of individual pulses from the pulsar B0329+54 at 111.4 MHz for July 26, 1999. An interval of values encompassing 180° is presented twice shifted by 180° with different signs. The three continuous curves for the polarization-angle variations correspond to the primary and secondary polarization modes in a rotating-vector model with $\alpha = 59°, \beta = -5°$, and $\varphi_0 = -1°$. The dashed curves show the boundaries separating the two orthogonal polarization modes. We carried out polarization measurements only for recordings with amplitudes exceeding 4σ, where σ is the rms scatter of the noise. Pulses with amplitudes below this threshold were not considered in the subsequent analysis; in a typical recording, such low-amplitude pulses comprised about 25% of the total number of pulses and about 2.5% of the integrated intensity.

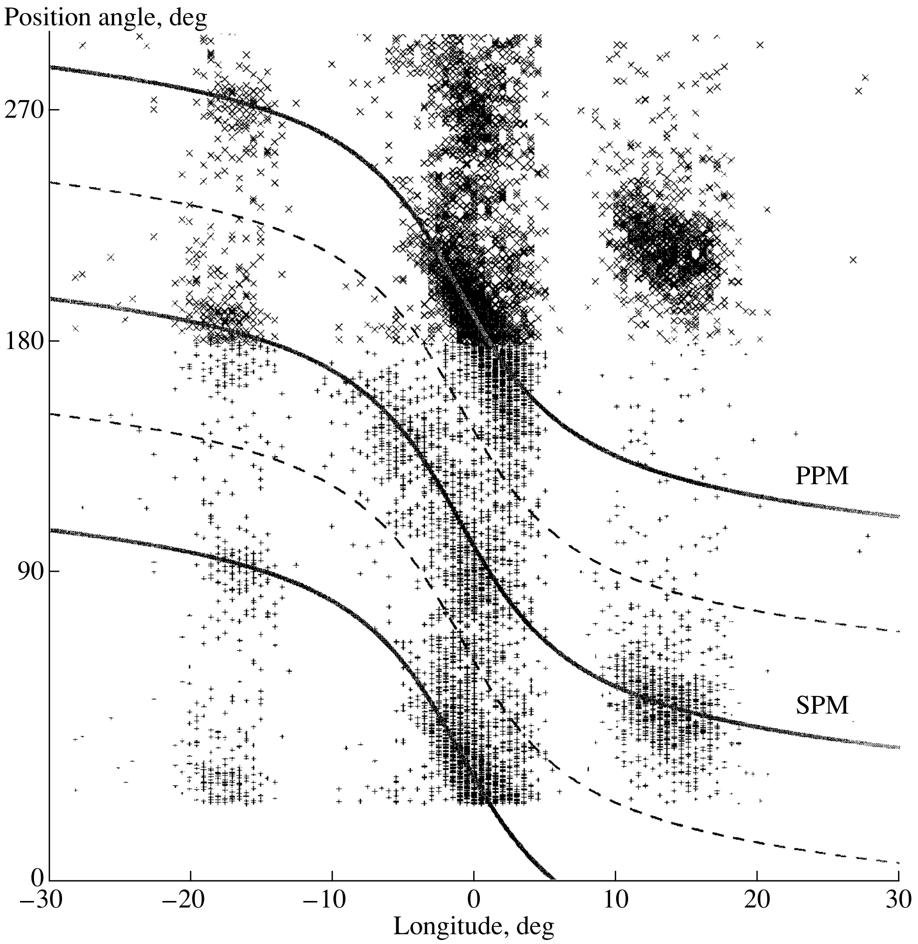

Fig. 7. Longitude distribution of the polarization position angles of individual pulses of B0329+54 at 111.4 MHz for the July 26, 1999, observing session. An interval of values encompassing 180° is presented twice shifted by 180° with different signs. The continuous curves for the polarization-angle variations correspond to the primary and secondary polarization modes in a model with a rotating polarization vector with parameters $\alpha = 59°, \beta = -5°$, and $\varphi_0 = -1°$. The dashed curves show the boundaries separating the two orthogonal polarization modes.

For each polarization mode, we derived the longitude distributions for the number of readings (n), average ($\langle A\rangle$) and total ($\langle A\rangle \times n$) reading amplitudes, and average degree of linear polarization. Figure 8 shows these distributions for the normal (left) and abnormal (right) modes. We can see that, in the abnormal mode, the number and average amplitude of readings corresponding to the secondary polariztion mode increased, leading to an increase in the total intensity for this polarization mode. In Fig. 8, the total amplitude for the SPM readings is normalized to the amplitude for the PPM; in the abnormal mode on July 24, 1999, it grew by a factor of three compared to the normal mode on July 23, 1999. Measurements of this sort collected for six days of observations (three in the normal mode and three in the abnormal mode) indicate that the relative intensity of the secondary polarization mode grew by a factor of 2.5 in the abnormal mode, leading to depolarization of the corresponding integrated pulses.

The profiles of the average DLP for the two polarization modes also differ: the polarization of the SPM is systematically higher in the outer components and lower in the central component.

Dependence of the Shape and Polarization of the Integrated Pulse on the Intensity of Individual Pulses

We noticed that, in the total-amplitude profile in Fig. 8, the outer components for the PPM are shifted toward the central longitude, compared to their positions for the SPM. In other words, the total opening angle of the radiation cone formed by the primary polarization mode is systematically smaller than that formed by the secondary polarization mode by several degrees. In addition, the average amplitude of the subpulses $\langle A\rangle$ is higher for the PPM than for

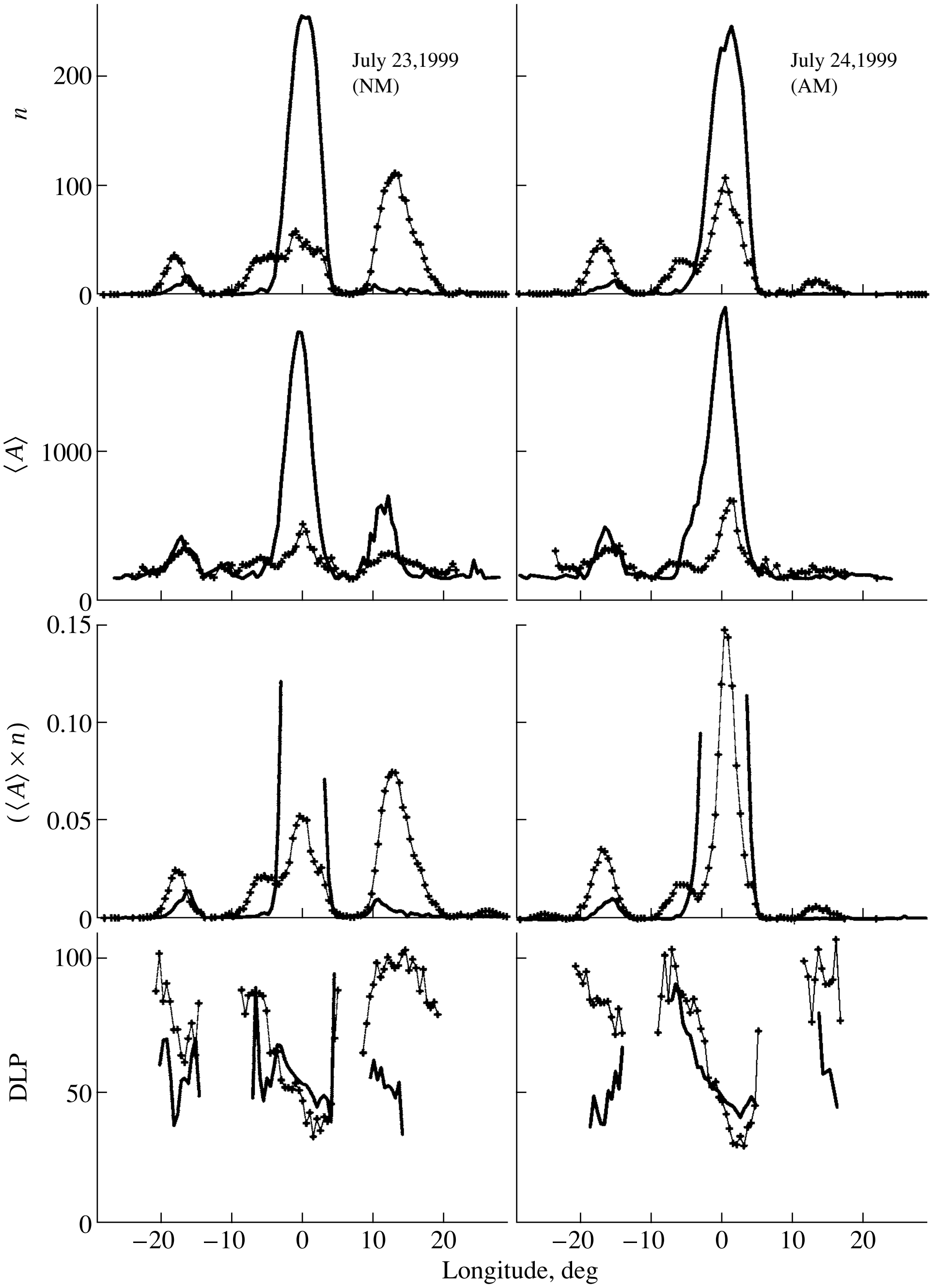

Fig. 8. Longitude distribution for four parameters characterizing individual pulses after dividing them into two groups corresponding to the primary (thick curves) and secondary (thin curves) polarization modes (PPM and SPM): the number of readings, average amplitude, total amplitude, and average degree of linear polariztion (from top to bottom). Observations of the normal mode on July 23, 1999, and of the abnormal mode on July 24, 1999, are shown to the left and right, respectively. The total amplitude for readings for the SPM is normalized to the amplitude for the PPM.

the SPM for all components of the intensity profile (except for component II). To test for a possible relationship between the relative intensity of the radiation and its polarization mode, we determined the polarization characteristics of the central components of pulses in groups with high and low intensities. Figure 9 shows examples of integrated pulses obtained by averaging individual pulses with amplitudes higher than 50σ (strong pulses) and lower than 10σ (weak pulses). The polarization position angle for the sum of strong pulses is determined by the primary polarization mode. This mode dominates in the central regions of individual strong pulses, while the secondary polarization mode dominates at their fronts.

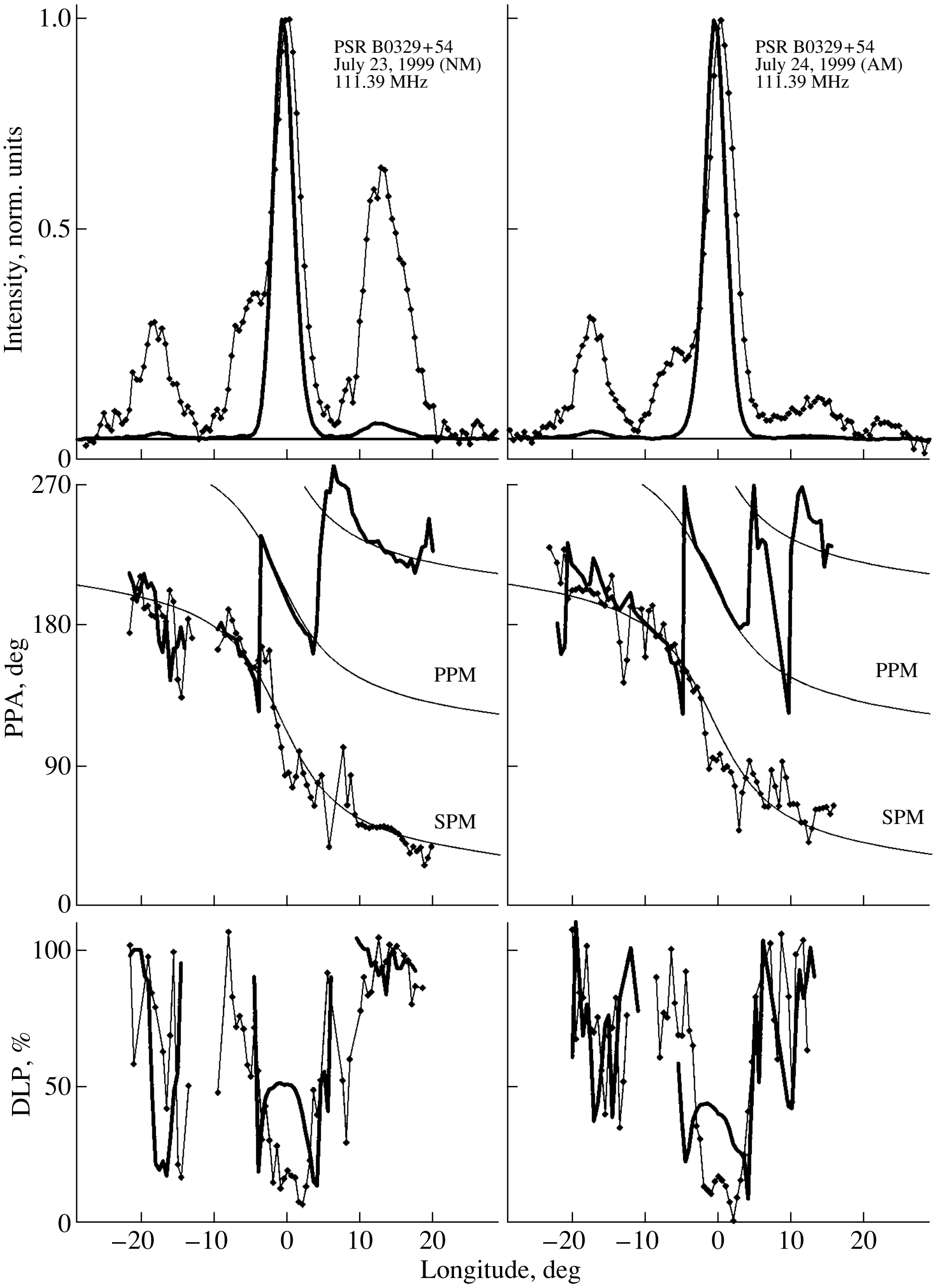

Fig. 9. Polarization characteristics of integrated pulses obtained by summing individual pulses with amplitudes higher than 50σ (thick curve, $N = 85$) and lower than 10σ (thin curve, $N = 220$). Observations of the normal mode on July 23, 1999, and of the abnormal mode on July 24, 1999, are shown on the left and right, respectively.

In this case, there is no significant depolarization at the center of the integrated pulse.

As expected, the run of the polarization position angle with longitude for the summed weak pulses is determined by the secondary polarization mode. The significant depolarization of the central component in this second group indicates that the weak pulses tend to belong entirely to one or the other of the two polarization modes. This gives rise to fluctuations in the PPA by 90° from pulse to pulse, and the integrated pulse at this longitude becomes appreciably depolarized.

When accumulating the strong pulses, the position of the maximum of the central component shifts to the left, toward earlier longitudes, and the relative amplitude of the outer components decreases. The total variation of the phase of the central component during the progressive decrease in the intensity of the summed pulses at 111.4 MHz was 1°, or 2 ms.

4. SUMMARY OF MAIN RESULTS

We have presented measurements of the characteristics of individual pulses of B0329+54 near times of switches from its normal to its abnormal mode at 111.4 MHz.

Currently, times of switches from one mode to the other have been determined only at frequencies higher than 400 MHz [1–3]. Times of mode switches at higher frequencies have been determined only to within the intervals used for preliminary averaging of the data, carried out to achieve adequate signal-to-noise ratios: 3 min (252 pulses) at 410 MHz [1] and 2695 MHz [2], and 2.4 and 3.9 min (200 and 330 pulses) at 1400 and 9000 MHz [3]. In our studies, we did not perform any preliminary averaging of short series of pulses.

The main results of our measurements at 111.4 MHz are listed below.

1. We have detected four events corresponding to switches from the normal to the abnormal mode in pulsar B0329+54 in four separate five-minute observing sessions. The probability of finding such events in our observations was 4%.

2. Variations in the strengths of the outer components near mode switches do not always occur simultaneously. The abnormal pulse shape is established as a result of both rapid (occuring over seconds) and slower (occuring over minutes) variations in the intensities and phases of subpulses. Times of mode switches at low frequencies are defined, first and foremost, by a decrease in the energy at the longitude of component IV by a factor of four to five and a jump-like variation in the position of the central component by approximately 1 ms (0.5°). Variations in the strength of component I can occur at the time of a mode switch (events E1, E3), preceeding a mode switch (E4), or be absent in a several-minute interval surrounding a mode switch (E2).

3. At 111.4 MHz, the magnitudes of the phase variations during mode switches are different for different components in the integrated profile; the first three components shift by 1.2 ms, while the phase of component IV remains constant.

4. In about half the cases, the rapid and simultaneous change in the shape and phase of the integrated profile associated with a mode switch can be preceeded by an additional smooth variation in the phases of individual pulses at the longitudes of component I and the central component over intervals of at least several minutes.

5. The degree of linear polarization of the central component in the abnormal mode is, on average, 8% lower than in the normal mode. This is due to an increase in the relative strength of the secondary polarization mode by a factor of two to three.

6. The shapes, phases, and polarizations of individual pulses of B0329+54 depend on their relative intensities.

5. DISCUSSION

Factors That Can Cause Variations in Pulse Shape

Let us consider several possible origins of the observed variations in the shape of the integrated profile.

A. Variations in the Relative Intensities of the Polarization Modes.

A progressive increase in the longitude of the central component of pulsar B0329+54 during a decrease in the intensity of summed pulses, accompanied by variations in the relative amplitudes of the outer and central components, was first discovered at the high frequency of 1415 MHz [9]. McKinnon and Hankins [9] suggested that this was due to random relocations of a hot spot whose radiation corresponded to the central component relative to the line of sight to the observer, with an amplitude of 130 m at the stellar surface, near the magnetic pole.

The correlation between the intensity of a pulse and its polarization mode we have detected suggests an alternative interpretation of this effect: the position and amplitude of a pulse are determined by which of the two polarization modes has made the dominant contribution to the radiation of this pulse.

When analyzing the dependences of the shapes, phases, and polarizations of pulses on their intensities for various threshold amplitudes, we found that the decrease in the degree of linear polarization of the central component and its delay by 1 ms can be obtained using recordings of normal-mode pulses simply by limiting the number of strong pulses that are included in the averaging. However, it is not possible to explain the phenomenon of mode switching purely as a consequence of a decrease in the number of strong pulses, since the amplitude of the central component should then decrease in the abnormal mode by a factor of two compared to the profile for the normal mode, which is not observed. In practice, no systematic decrease in the amplitude of this component in the abnormal mode is observed. In addition, when the amplitude of the central component decreases, the relative amplitude of component IV grows (Fig. 9), not decreases, as in the abnormal mode. Nevertheless, an increase of the relative strength of the secondary polarization mode in the central component in the abnormal mode could provide an additional shift of its phase toward higher longitudes.

The results presented in Fig. 8 show that in the outer components I and IV subpulses in which the primary polarization mode dominates appear more rarely; they are more intense and shifted toward the central longitude. This is confirmed by our observations of B0329+54 at 103 MHz [10]. In contrast to the situation for the central component, the intensity–phase correlation for component I can strengthen the longitude shift associated with the mode switches. The increase in the intensity of component I in the abnormal mode by a factor of 1.5 reflects an increase in the number of strong pulses, which are associated with the primary polarization mode, and are consequently shifted toward zero longitude. This could lead to a phase delay of component I in the abnormal mode. In contrast to component I, the amplitude and phase at the longitude of component IV are fully determined by the secondary polarization mode. This can clearly be seen in the distribution of polarization position angles in Fig. 7. The influence of the primary polarization mode on the position of this component in switches to the abnormal mode is negligible, and the phase of component IV remains constant.

Thus, although temporal variations in the efficiency of radiation in the two polarization modes cannot fully explain all the observed variations occuring during mode switches, they undoubtedly play an important role in the formation of the two modes.

B. Variations in the Height of the Emitting Region above the Stellar Surface.

Bartel *et al.* [3] suggest that the longitudinal shift and reduction in the total pulse width are due to the fact that that the abnormal-mode radiation originates from a lower level in the pulsar magnetosphere, where the divergence of the dipolar magnetic-field lines is smaller. This model can be used to estimate the difference in the heights of the regions in which the two orthogonal polarization modes are radiated. The distance between the outer components in the profile is smaller for the primary polarization mode than for the secondary polarization mode (Fig. 8); this could indicate that the PPM is radiated at a lower height above the stellar surface. The heights at which the two orthogonal polarization modes are radiated can be estimated from the distances $\Delta\varphi$ between the outer components for the two modes in Fig. 8. At 111.4 MHz, these distances are 26.6° and 30.3° for the PPM and SPM, respectively. In the case of an aligned rotator, the distance between the outermost conal components is determined by some of the most peripheral of the open magnetic field lines, and the height of the region of radiation r can be derived from the well known relation $\sin^2(\Delta\varphi/2) = r/r_{LC}$, where r_{LC} is the radius of the light cylinder.

For an oblique rotator (for B0329+54, the angle between the rotational axis and magnetic-dipole axis α is approximately 60°), we must introduce a parameter $f(\alpha)$ into the above expression to account for the dependence of the opening angle of the cone of magnetic field lines on the angle α. It can be shown that $f(\alpha)$ varies from 1 to 0.6 when α varies from 0 to 90°; $f(\alpha) = 0.7$ when $\alpha \cong 60°$. For B0329+54, the influence of β on the angular distance between the components can be neglected, since this angle is close to zero ($\beta = -5°$) and the line of sight passes nearly through the center of the cone. Using the final expression $r = (cP/2\pi)\sin^2(\Delta\varphi/2)/f^2(\alpha)$, we obtain $r(\mathrm{PPM}) = 3.7 \times 10^8$ cm and $r(\mathrm{SPM}) = 5 \times 10^8$ cm. The corresponding difference in the heights for the radiation in the two orthogonal polarization modes is $\Delta r = 1.3 \times 10^8$ cm. This can be taken to correspond to the extent of the radiating region in height at frequencies below 100 MHz.

The difference in the heights at which the orthogonal modes are radiated was estimated earlier for another pulsar, B0834+06, at 1404 MHz [11], based on the phase shift of pulses formed in the two polarization modes. This analysis used a formula for the time delays of pulses radiated at different distances from the stellar surface: $\Delta r = \Delta\varphi cP/360° = 1.3 \times 10^8$ cm. The resulting value is similar to our estimate of this height difference for B0329+54.

Due to this height difference, random time variations in the efficiency of radiation of the two orthogonal modes could lead to longitude shifts in the resulting observed pulse.

C. Variations in the Subpulse Drifting Behavior

In our opinion, there is another mechanism that could lead to longitude shifts of the pulses, namely subpulse drift.

Variations in the shape of the integrated pulse profile accompanied by a strengthening of the secondary polarization mode were recently discovered for another mode-switching pulsar, B0943+10, at 103 and 430 MHz [12]. The subpulses of B0943+10 demonstrate a stable drift in the normal mode, which is disrupted in the abnormal mode. Variations in the drift behavior of subpulses in the abnormal mode accompanied by phase shifts in at least one component have also been observed for other mode-switching pulsars [3]. We believe that we can draw an analogy with B0329+54 and relate the observed variations in the shape of an integrated pulse, its position, and its polarization with variations in the normal subpulse drift pattern. In spite of the fact that no clear subpulse drift pattern has been identified for B0329+54, there are indirect indications that such a pattern is present in this mode-switching pulsar as well. For example, variations in the character of subpulse amplitude fluctuations have been detected during mode switches, especially near the central component [3]. In addition, there is evidence for the presence of fine structure within each component in the individual pulses recorded by us at 111.4 MHz. Each element of this structure has a full width at half maximum of about 5 ms; there from one to three elements for each of various pulses, with the total component width varying from 5 to 12 ms. Occasionally, two maxima separated by 3 to 8 ms, depending on the relative intensities of the subcomponents, are observed at the longitude of the central component. These relative intensities may be related to the drift phase of each element of the subcomponent structure.

Thus, there exists an entire complex of factors capable of bringing about the observed variations in the shapes and polarizations of the integrated pulses during mode switches. These factors may be activated by variations in the physical conditions at the stellar surface in the region of generation of the primary relativisitic particles.

Slow Component Establishing the Abnormal Mode

A slow component in the variations of pulsar radiation characteristics preceding mode switch in some pulsars is a complementary factor in establishing abnormal mode. It was first detected in pulsar B0355+54 [13]. More specifically, a slow growth in the degree of linear polarization of one component of the integrated pulse was discovered. This pulsar remained unique in this respect for a long time, until a slow (acting over tens of minutes) decay of one component lasting until a flare in the emission of another component was discovered in another mode-switching pulsar, B0943+10 [12]. We have now observed a similar phenomenon in B0329+54 at 111.4 MHz: slow variations in the phases of individual pulses lasting until a mode switch. Slow variations in the relative intensities of components of B0329+54 and the distances between them occuring over the course of an hour between successive mode switches are clearly visible at 1.4 and 9.0 GHz (Fig. 2) [3], a fact that was not commented on by Bartel *et al.* [3].

The absence of variations in the amplitude and phase of component I in event E2 during the first 2.5 min after the mode switch also suggests the existence of slow variations in the pulse characteristics. The standard profile for the abnormal mode of B0329+54 probably becomes firmly established only within several tens of minutes after the actual mode switch.

This leads us to conclude that the presence of long-term variations in the properties of the pulsar radiation associated with mode switches are an omnipresent property of all mode-switching pulsars and that it is worthwhile to continue searches for such variations in other pulsars of this type.

6. ACKNOWLEDGMENTS

The author thanks I.F. Malov and Prof. W. Sieber (Germany) for useful comments made during discussions of these results. This work was supported by the Russian Foundation for Basic Research (project no. 00-02-17447). The Large Phased Array of the Lebedev Physical Institute is supported by the Ministry of Industrial Science of the Russian Federation in the framework of the program "Unique Installations and Facilities of Russia" (reg. no. 01-11).

REFERENCES

1. A. G. Lyne, Mon. Not. R. Astron. Soc. **153**, 27 (1971).
2. K. H. Hesse, Astron. Astrophys. **27**, 373 (1973).
3. N. Bartel, D. Morris, W. Sieber, and T. Hankins, Astrophys. J. **258**, 776 (1982).
4. S. A. Suleĭmanova, Yu. V. Volodin, and Yu. P. Shitov, Astron. Zh. **65**, 349 (1988) [Sov. Astron. **32**, 177 (1988)].
5. S. A. Suleĭmanova, in *Pulsars* (Nova Science, New York, 1992), Vol. 196, pp. 55–89.
6. T. V. Shabanova, Astron. Zh. **71**, 132 (1994) [Astron. Rep. **38**, 115 (1994)].
7. V. Radhakrishnan and D. J. Cook, Astrophys. Lett. **3**, 225 (1969).
8. J. A. Gil and A. G. Lyne, Mon. Not. R. Astron. Soc. **276**, L55 (1995).
9. M. M. McKinnon and T. H. Hankins, in *The Magnetospheric Structure and Emission Mechanisms of Radio Pulsars (IAU Colloquium 128),* Ed. by T. H. Hankins, J. M. Rankin, and J. A. Gil (Padagogical Univ. Press, 1992), p. 190.
10. S. A. Suleĭmanova and V. D. Pugachov, Astron. Zh. **75**, 287 (1998) [Astron. Rep. **42**, 252 (1998)].
11. J. A. Gil, Astrophys. J. **314**, 629 (1987).
12. S. A. Suleĭmanova, V. A. Izvekova, J. M. Rankin, and N. Rathnashree, J. Astrophys. Astron. **19**, 1 (1998).
13. D. Morris, W. Sieber, D. C. Ferguson, and N. Bartel, Astron. Astrophys. **84**, 260 (1980).

Translated by D. Gabuzda

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 49–56. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 54–62.
Original Russian Text Copyright © 2002 by Slysh, Kalenskiĭ, Val'tts.

Methanol Radio Emission at Millimeter Wavelengths: New Masers at 1.3 and 2.8 Millimeters

V. I. Slysh, S. V. Kalenskiĭ, and I. E. Val'tts

Astro Space Center, Lebedev Institute of Physics, Russian Academy of Sciences, Profsoyuznaya ul. 84/32, Moscow, 117810 Russia

Received October 26, 2000

Abstract—The results of a search for maser emission in the methanol lines 8_{-1}–7_0E at 229.8 GHz, 3_{-2}–$4_{-1}E$ at 230.0 GHz, 0_0–$1_{-1}E$ at 108.9 GHz, and in the $J_1 - J_0E$ series near 165 GHz in star-forming regions are reported. At least two masers and two candidates have been detected at 229.8 GHz. Thus, methanol masers have been detected in the 1-mm band for the first time. At 108.9 GHz, masers have been detected toward G345.01+1.79 and possibly toward M8E as well. Thermal emission was found toward 28 objects. The 229.8-GHz sources are class I masers, whereas the 108.9-GHz sources are class II masers. An analysis using a large velocity-gradient method shows that the 229.8-GHz masers can appear at densities of about 3×10^4 cm^{-3}. The ratios of the flux densities in different class I lines toward DR 21(OH) and DR 21 West can be approximated in models with gas kinetic temperatures of about 50 K. Detection of the 108.9 GHz masers toward G345.01+1.79 and M8E may provide information about the geometry of these objects. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The methanol molecule (CH_3OH) is a slightly asymmetric top with hindered internal rotation; it possesses a multitude of allowed transitions at radio wavelengths. The radio emission of interstellar methanol has been actively studied since its discovery by Ball *et al.* [1] in 1970. Narrow, bright, undoubtedly maser lines have been detected at many frequencies, often superimposed on broader thermal features. Currently, methanol masers have been found in hundreds of star-forming regions in our Galaxy and the Large Magellanic Cloud [2]. A classification for methanol masers was first proposed by Batrla *et al.* [3] and later modified by Menten [4]. Menten [4] divides all methanol masers into two classes. Class I includes masers emitting in the lines 7_0–6_1A^+, 4_{-1}–3_0E, 8_0–7_1A^+, etc.; class II includes masers emitting in the lines 2_{-1}–3_0E, 5_1–6_0A^+, 3_1–4_0A^+, etc. A number of theoretical studies devoted to methanol excitation have shown that class I masers are pumped by collisional excitation of rotational levels, and class II masers, by radiative transitions under the action of strong radiation [5, 6]. As a rule, methanol masers simultaneously emit several lines belonging to the same class. It is currently believed that lines belonging to different classes cannot be inverted simultaneously. However, some class I and class II maser sources are associated with the same star-forming regions, and are so closely spaced on the celestial sphere that it is virtually impossible to distinguish between their positions in single-dish observations.

The strongest class I masers have been detected in the 7_0–6_1A^+ transition at 7 mm and the 4_{-1}–3_0E transition at 8 mm. At shorter wavelengths, in the 3-mm and 2-mm bands, weaker masers in the 5_{-1}–4_0E, 6_{-1}–5_0E, and 8_0–7_1A^+ lines have been observed. The situation is similar for class II masers: the strongest masers in the 2_{-1}–3_0E and 5_1–6_0A^+ lines are observed at centimeter wavelengths, and weaker ones are observed in the 3_0–4_1A^+ and J_0–$J_{-1}E$ lines in the 3-mm and 2-mm bands. This wavelength dependence of the line intensity is explained by the decrease in the degree of inversion and optical depth with decreasing wavelength. When deriving the parameters of a maser source, it is useful to know the limiting wavelengths at which maser emission is possible. Therefore, observations of methanol masers at wavelengths shorter than those observed earlier are of interest.

We conducted a search for masers in the 8_{-1}–7_0 line of E-methanol at 229.8 GHz (1 mm). Previously, no searches for methanol masers had been carried out at such high frequencies. According to the theoretical models of Sobolev *et al.* [6], 8_{-1}–7_0E is a class I transition, and 229.8-GHz masers should be observed together with masers in the 7_0–6_1A^+ and 4_{-1}–3_0E transitions, as well as other transitions of this class. In addition to the 229.8-GHz line, the

1063-7729/02/4601-0049$22.00 © 2002 MAIK "Nauka/Interperiodica"

Table 1. Parameters of the observed methanol lines

Transition	Frequency, MHz	Line strength*	E_u/k, K**
$0_0.1_{-1}E$	108894.20	0.4713	5.23
$1_1{-}1_0E$	165050.19	0.6503	15.48
$2_1{-}2_0E$	165061.14	1.0802	20.13
$3_1{-}3_0E$	165099.31	1.5045	27.10
$4_1{-}4_0E$	165190.53	1.9207	36.40
$5_1{-}5_0E$	165369.44	2.3258	48.01
$8_{-1}{-}7_0E$	229758.76	2.4312	81.27
$3_{-2}.4_{-1}E$	230027.06	0.3543	31.96

Note: *In accordance with Cragg *et al.* [8], **energy is measured relative to the E-methanol ground level $1_{-1}E$.

$0_0{-}1_{-1}E$ line at 108.9 GHz (3 mm), a series of $J_1{-}J_0E$ lines at 165 GHz (2 mm) and the $3_{-2}{-}4_{-1}E$ line (1 mm) were observed. According to statistical equilibrium calculations [6, 7], these lines can be inverted in class II sources.

2. OBSERVATIONS

The observations were carried out on the 30-m radio telescope of the Institut de la Radioastronomie Millimétrique (IRAM) in Pico Veleta (Spain) on August 29–31, 1995. The source list included 36 objects in which strong class I and II masers had been detected earlier in other methanol lines. Table 1 lists the frequencies and strengths of the lines observed. Table 2 presents the antenna parameters: antenna efficiency (AE), main-beam efficiency ($B_{\rm eff}$), forward hemisphere scattering factor ($F_{\rm eff}$), halfpower beamwidth (HPBW), and antenna temperature to flux density conversion factor $S_\nu/T^*_{\rm A}$ at 108, 165, and 229 GHz.

The observations were conducted in an on–off mode with three receivers in all three bands simultaneously. We performed the calibration using a chopper-wheel method. The noise temperatures of the receivers at 1, 2, and 3 mm were $\sim$ 100 K, and the system noise temperature varied in the range 300–1500 K at 3 mm and 1000–3000 K at 1 and 2 mm[1], depending on the weather and source elevation above the horizon. The pointing accuracy was checked using observations of continuum sources, and was no worse than 3$''$.5. For several sources, we obtained maps with a small number of points (5–10).

We used a 512-channel spectrometer with a resolution of 1 MHz (1.3 km/s) for the observations at 1 mm. The analysis band enabled us to observe the frequencies of the $8_{-1}{-}7_0E$ and $3_{-2}{-}4_{-1}E$ lines simultaneously. The band also included the $19_5{-}20_4A^+$ and $19_5{-}20_4A^-$ methanol lines. To observe the $8_{-1}{-}7_0E$ and $3_{-2}{-}4_{-1}E$ lines with high frequency resolution, the receiver backend was connected in parallel to a 256-channel spectrometer with a resolution of 100 kHz (0.13 km/s), split into two sections of 128 channels each.

Table 2. Parameters of the IRAM 30-m radio telescope

Frequency, GHz	AE	$B_{\rm eff}$	$F_{\rm eff}$	HPBW, arcsec	$S_\nu/T^*_{\rm A}$, Jy/K
108	0.57	0.68	0.92	22.0	6.3
165	0.41	0.50	0.90	15.0	8.6
229	0.32	0.39	0.86	10.4	10.5

We used another 512-channel filter-bank spectrometer with a resolution of 1 MHz (1.8 km/s) for the 2 mm observations. The analysis band enabled us to observe the $J = 1{-}4$ lines of the $J_1{-}J_0$ series simultaneously. A section of an autocorrelator (1000 channels) with a resolution of 80 kHz (0.14 km/s) was connected in parallel to enable observation of the $J = 1{-}3$ lines with high resolution.

In four sources, we observed at the frequency of the $5_1{-}5_0E$ line by retuning the 2-mm receiver to this frequency.

For the 3-mm observations, we used the second section of the autocorrelator, with a resolution of 40 kHz (0.11 km/s).

The data were processed with the CLASS software package. When fitting Gaussians to the $J_1{-}J_0E$ lines, we assumed their LSR velocities and linewidths to be identical.

3. RESULTS

We observed 36 star-forming regions and one galaxy. Emission in at least one methanol line was detected in 30 sources.

In the $8_{-1}{-}7_0E$ transition, we detected narrow maser lines toward DR 21(OH) and DR21 West. These are the first detections of methanol masers at 1 mm. Broad, quasi-thermal lines were detected toward 18 sources. We believe that at least two of these lines represent blends of narrow maser features (see the next section). We detected maser emission in the $0_0{-}1_{-1}E$ line toward G345.01+1.79 and possibly M8E. In addition, we have detected 28 thermal sources in this line; among these was the galaxy IC 342, from which we detected emission at the sensitivity limit. Only thermal lines were found in the $3_{-2}{-}4_{-1}E$ and $J_1{-}J_0E$ transitions. For sources in

---

[1] Some sources were observed with higher noise temperatures. Such cases are marked in Table 3 by asterisks.

Table 3. Gaussian parameters of the lines

Source	RA (1950) Dec (1950)	Transition	$\int T_A^* dV$, (K km)/s	$V_{\rm LSR}$ km/s	ΔV km/s	T_A^*, K
G345.01+1.79	$16^{n}53^{m}19.6^{s}$	I	5.9(1.0)*	−13.71(0.47)	4.91(0.83)	
	−40°09′46″	II				<0.90
		III	7.2(1.7); 13.7(1.7); 7.8(1.7); 8.7(1.7); 3.3(1.6)*	−13.95(0.32)	6.01(0.00)	
		IV	0.7(0.2)	−22.55(0.11)	1.03(0.25)	
		IV	0.8(0.1)	−21.90(0.01)	0.49(0.04)	
		IV	2.2(0.1)	−13.60(0.09)	4.00(0.22)	
NGC 6334I(N)	$17^{n}17^{m}35.4^{s}$	I	10.4(0.8)*	−4.88(0.20)	5.09(0.47)	
	−35°42′23″	II				<0.6*
		III	7.3(0.9); 8.7(0.9); 9.3(0.9); 7.0(0.8); $8.1(1.4)^3$	−4.27(0.16)	5.34(0.22)	
		IV	11.2(0.2)	−4.64(0.04)	5.42(0.10)	
M8E	$18^{n}01^{m}49.7^{s}$	I				<0.27
	−24°26′56″	II				<0.27
		III				<0.45
		IV	0.4(0.05)	10.66(0.04)	0.83(0.13)	0.45
L379 IRS3	$18^{n}26^{m}32.9^{s}$	I	5.4(0.5)	20.45(0.22)	4.72(1.07)	
	−15°17′20″	II				<0.45
		III	1.9(0.7); 3.2(0.7); 1.9(0.7); 3.1(0.7)	19.78(0.34)	4.56(0.74)	
		IV	3.5(0.2)	19.32(0.13)	5.56(0.43)	
DR 21 West	$20^{n}37^{m}07.8^{s}$	I	1.1(0.2)	−2.55(0.02)	0.34(0.04)	
	42°08′45″	I	2.7(0.3)	−2.16(0.12)	1.70(0.25)	
		II				<0.25
		III				<0.4?
		IV				<0.2
DR 21(OH)	$20^{n}37^{m}12.3^{s}$	I	3.1(0.2)	−1.64(0.14)	4.18(0.28)	
	42°12′11″	I	2.0(0.1)	0.29(0.01)	0.65(0.03)	
		II				<0.2
		III	1.7(0.3); 2.3(0.3); 2.2(0.3); 2.5(0.3)	−2.65(0.20)	5.76(0.31)	
		IV	2.1(0.1)	−2.75(0.08)	4.71(0.17)	

Note: Transition designations: I—8_{-1}–7_0E, II—3_{-2}–$4_{-1}E$, III—J_1–J_0E, IV—0_0–$1_{-1}E$. The last column lists 3σ upper limits for the antenna temperature. An asterisk means that the system noise temperature was higher than 3000 K during the observations.

which masers were detected in at least one transition, the line parameters are listed in Table 3 and the spectra are presented in Fig. 1. The results of the observations for other sources are given in a separate paper [18].

4. COMMENTS ON INDIVIDUAL SOURCES

In this section, we briefly describe the emission detected. A more complete description of the maser sources G345.01+1.79 and M8E is given in [9, 10].

G345.01+1.79. In this object, we detected a line at 108.9 GHz that can be decomposed into two components. A narrow line at −22 km/s coincides

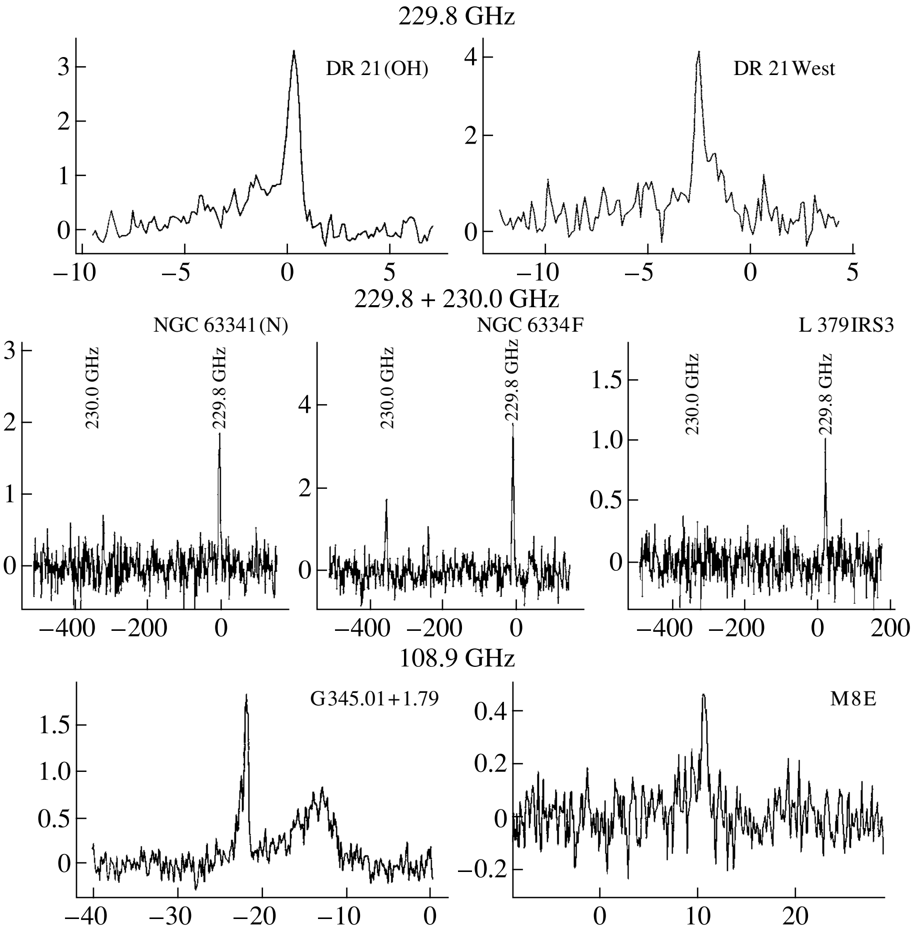

Fig. 1. Spectra of sources in which 229.8-or 108.9-GHz maser emission has been detected; the spectrum of the thermal source NGC 6334F with the 8_{-1}–7_0E (229.8 GHz) and 3_{-2}–$2_{-1}E$ (230.0 GHz) lines is given for comparison. Radial velocities (km s^{-1} relative to the local standard of rest of 8_{-1}–7_0E lines (upper and middle lines) are plotted along the X axis. System temperatures (in Kelvins) are plotted along the Y axis.

in velocity with the masers at 6.7, 12.2, 107.0, and 157 GHz [11, 13, 14] and is undoubtedly a maser line. G345.01+1.79 is currently the only source in which a 108.9-GHz maser has been confidently detected. In addition to the maser line, we detected a broad thermal line at this frequency.

M8E. Emission was detected only at 108.9 GHz. The strongest line was recorded at the point (−10″, 0″) at a projected distance of 12″ to the southeast of the 44.1-GHz maser, which is one of the strongest class I masers. The linewidth is small, 1–2 km/s, depending on the position. This suggests that we are observing either cool, quiescent gas or maser emission at 108.9 GHz. Masers in the 0_0–$1_{-1}E$ line belong to class II [13]. Since none of the strongest class II masers (5_1–6_0A^+ at 6.7 GHz) have been detected in M8E [11], quiescent-gas emission seems more probable. However, we cannot rule out the possibility that, in this source, we are observing a maser pumped by the radiation of an HII region with a relatively low emission measure. This radiation may not be strong enough to populate the 5_1A^+ level. The presence of OH masers [12], which usually accompany class II methanol masers, is consistent with this possibility. A detection of emission in the 2_0–$3_{-1}E$ line at 12.2 GHz would confirm this hypothesis.

L379 IRS3 and NGC 6334I(N). We detected strong, broad lines at 229.8 GHz. Lines found earlier at the same radial velocities in other class I transitions are blends of maser features [15–17]. It is quite probable that the 229.8-GHz lines are also maser lines and consist of narrow features. The nonthermal character of the emission in the 8_{-1}–7_0A^+ line is confirmed by the ratio of its intensity to the intensity in the 3_{-2}–$4_{-1}E$ line at the nearby frequency of 230.0 GHz. In sources with obviously thermal 229.8-GHz emission lines, such as NGC 6334F, Ori KL, G34.26+0.15, W51E1/E2, and W31(1), these intensity ratios are 1.8–3 [18], while, in L379 IRS3 and NGC 6334I(N), the 3_{-2}–$4_{-1}E$ line was not detected

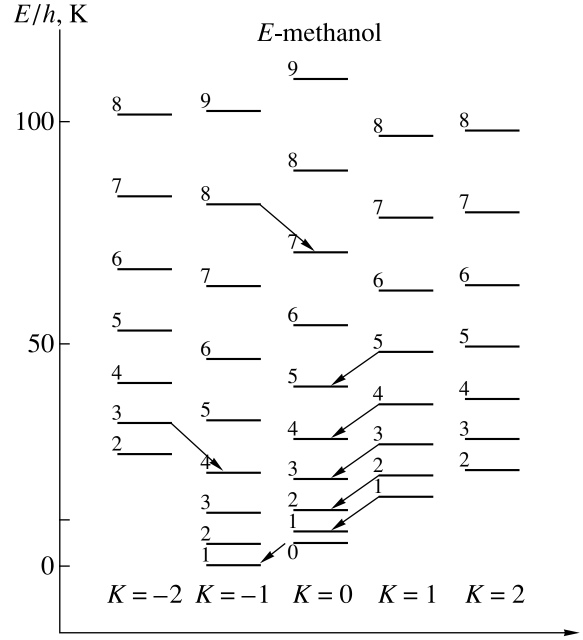

Fig. 2. Energy-level diagram of methanol. The arrows show the transitions detected.

at all, and the corresponding intensity ratio exceeds 2.5 in L379 IRS3 and 3.4 in NGC 6334I(N) (Table 3).

DR 21West. We detected a narrow 229.8 GHz line superimposed on a weaker and broader feature. The narrow line coincides in radial velocity with a strong maser feature in the 7_0–$6_1 A^+$ line at 44.1 GHz [16] and is undoubtedly a maser line. It is interesting that the linewidth at 229.8 GHz is smaller than at 44.1 GHz. This may indicate a saturation of the 7_0–$6_1 A^+$ line, which results in its broadening.

DR 21(OH). We detected a narrow undoubtedly maser line at 229.8 GHz at (0″, 0″), whose radial velocity coincides with that of the strongest maser features in the 7_0–$6_1 A^+$ line at 44.1 GHz [16]. It is possible that the broader feature at (0″, 0″) and the broad asymmetric feature at (10″, 0″) we have detected at 229.8 GHz are blends of the 44.1 GHz maser features.

5. DISCUSSION

5.1. Excitation of Methanol Masers

The upper signal levels of most class I maser transitions lie on the main ladder; i.e., on the sequence of levels with the same quantum number K that contains the ground level. In the case of E-methanol, the ladder of levels with $K = -1$ is the main ladder (Fig. 2). On the contrary, the upper levels of class II maser transitions lie on side ladders. Statistical equilibrium calculations (Fig. 3a) show that, in the case of collisional excitation, the main ladder is overpopulated relative to the adjacent ladders, and transitions with their upper level on the main ladder and lower level on the $K = 0$ side ladder, such as 8_{-1}–$7_0 E$, are inverted.

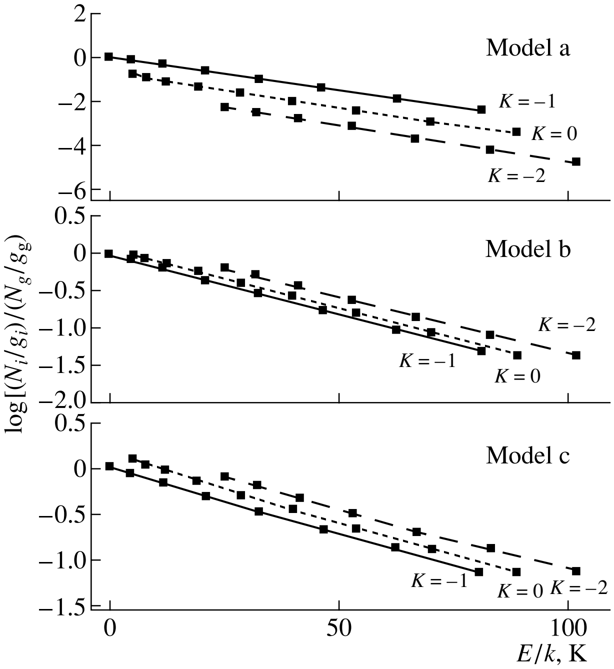

Fig. 3. Ratios of the E-methanol level populations to the population of the ground level $1_{-1}E$, obtained by LVG statistical equilibrium calculations. Points corresponding to the main ladder ($K = -1$) are connected by a solid line; points corresponding to the $K = 0$ and $K = -2$ ladders are connected by the dotted and dashed lines. Models a, b, and c are given in Table 4.

An external radiation field modifies this pattern. Under the action of radiation, the side-ladder levels are populated more efficiently (Fig. 3, models b, c). As soon as a certain radiation intensity is achieved, the inversion of the class I transitions disappears, but instead the class II transitions are inverted. This is illustrated by the data of Fig. 3 and Table 4, which present the excitation temperatures of four class I and four class II transitions calculated using the large-velocity-gradient (LVG) method for three models. In the absence of external radiation (model a), the class I transitions are inverted. When the radiation temperature exceeds the kinetic temperature of the gas, the inversion of class I transitions disappears, but instead the class II 2_0–$3_{-1}E$ transition is inverted (model b). In this model, the shorter-wavelength 0_0–$1_{-1}E$, 4_1–4_0E, and 3_{-2}–$4_{-1}E$ class II transitions remain noninverted. However, upon further increase of the external radiation intensity (model c), these transitions are also inverted.

Table 4 also demonstrates reductions of the degree of inversion and line optical depth with increasing

Table 4. Results of statistical equilibrium calculations for class I methanol masers and transitions corresponding to class II

Class	Transition	Frequency, GHz	Excitation temperature and optical depth					
			model a		model b		model c	
			$T_{\rm ex}$, K	τ	$T_{\rm ex}$, K	τ	$T_{\rm ex}$, K	τ
I	$4_{-1}-3_0E$	36.169	−1.1	−1.2	9.0	0.8	4.3	0.9
I	$5_{-1}-4_0E$	84.521	−2.8	−0.6	14.0	0.9	8.8	0.9
I	$6_{-1}-5_0E$	132.891	−4.9	−0.2	16.4	0.8	12.6	0.7
I	$8_{-1}-7_0E$	229.759	−10.6	−0.01	19.5	0.4	19.5	0.4
II	4_1-4_0E	165.190	2.5	0.3	48	0.8	−136	−0.2
II	$2_0-3_{-1}E$	12.179	0.4	1.8	−6.0	−0.5	−1.7	−0.9
II	$0_0-1_{-1}E$	108.894	2.8	1.6	60	0.3	−25	−0.4
II	$3_{-2}-4_{-1}E$	230.027	2.4	0.5	157	0.08	−42	−0.2

Note: In model a, the temperature of the external radiation is 2.7 K, the density is 10^5 cm^{-3}, and the specific methanol column density is 10^{-4} cm^{-3} (km/s pc^{-1})$^{-1}$. In model b, these quantities are 50 K, 6×10^{15} cm^2, and 10^{-3} cm^2 (km/s pc^{-1})$^{-1}$. In model c, they are 150 K, 10^{15} cm^2, and 6×10^{-3} cm^{-3} (km/s)$^{-1}$. In all models, the gas kinetic temperature is 25 K. The level population ratios are characterized by rational tempratures of 13, 25, and 26 K in models a, b, and c, respectively.

Table 5. Observed and model flux density in the class I maser lines

Transition	Flux density, Jy					
	DR 21(OH)		DR 21 West		W33 Met	
	model	observations	model	observations	model	observations
$4_{-1}-3_0E$	15.0	13 [20]	50.0	35 [20]	6.0	11 [20]
$5_{-1}-4_0E$	37.8	75[1]	82.0	110 [4]	9.0	18 [21]
$6_{-1}-5_0E$	45.6	67 [15]	82.5	64 [15]	10.7	15 [15]
$8_{-1}-7_0E$	32.7	25	40.0	34	9.7	<5
$9_{-1}-8_{-2}E$	0.003	<0.23 [22]	0.004	<0.25[22]	0.002	0.80 [22]
$7_0-6_1A^+$	35.9	340 [17]	52.5	240 [17]	20.2	70 [17]
$8_0-7_1A^+$	27.2	175[1]	28.5	150 [4]	11.7	13 [23]

Note: The model kinetic temperatures, densities, and specific methanol column density is were DR 21(OH): 60 K, 5.6×10^4 cm^{-3}, 4×10^{16} cm^2 (km/s)$^{-1}$; DR 21 West: 55 K, 3.2×10^4 cm^{-3}, 4×10^{-3} cm^{-3} (km/s)$^{-1}$; W33 Met: 55 K, 5.6×10^4 cm^{-3}, 1.2×10^{17} cm^2 (km/s)$^{-1}$. The observed fluxes without references were obtained in the present work. The gas kinetic temperature was 25 K in all models.

frequency. For instance, in Model a, the optical depth and excitation temperature of the lowest-frequency class I maser considered, $4_{-1}-3_0E$, are −1.2 K and −1.1 K, while those of the highest-frequency maser, $8_{-1}-7_0E$, are −0.01 K and −10.6 K (larger absolute values of the temperature mean smaller degrees of inversion).

5.2. Class I Masers

We have detected narrow maser features in the $8_{-1}-7_0E$ line toward DR 21(OH) and DR 21 West. In other sources, we found only broad quasi-thermal lines. However, at least some of these lines may consist of narrow maser features (see comments for L379 IRS3 and NGC 6334I(N)).

To estimate the conditions under which the masers in the $8_{-1}-7_0E$ line are generated, we analyzed the results of LVG statistical equilibrium calculations. The temperature was varied from 10–100 K[2] in steps

---

[2] In addition, we performed modeling for a kinetic temperature of 130 K.

Table 6. Modeling of class II maser excitation under the action of emission from a compact HII region ("HII") and hot dust ("dust"). For comparison, observed fluxes for G345.01+1.79 are given

Transition	Frequency, GHz	Observed flux, Jy	Model flux, Jy	
			HII	dust
$5_1{-}6_0A^+$	6.7	508 [11]	508	508
$2_0{-}3_{-1}E$	12.2	310 [11]	289	341
$3_1{-}4_0A^+$	107.0	85.5 [13]	171	166
$0_0{-}1_{-1}E$	108.9	9.5	89	22
$2_1{-}3_0A^+$	156.6	21.4 [14]	88	92
$4_0{-}4_{-1}E$	157.2	54.9 [14]	72	23
$4_1{-}4_0E$	165.2	<7	65	<0.01
$3_{-2}{-}4_{-1}E$	230.0	<9	185	–

Note: The model fluxes were calculated using the data from [6]. Observed fluxes given without references were obtained in the present work. The observed flux for the $4_1{-}4_0E$ transition is given at the radial velocity of the maser line (−22.5 km/s).

of 5 K, the density was varied from 10^4–10^8 cm^{-3}, and the specific methanol column density varied in the range 4×10^{13}–4×10^{17} (km/s)$^{-1}$. We assumed the absence of any external radiation, except for the cosmic microwave background.

The calculations show that masers in the $8_{-1}{-}7_0E$ line that are comparable in intensity with masers in $4_{-1}{-}3_0E$, $5_{-1}{-}4_0E$, and other class I lines can arise at gas temperatures of 50–60 K, typical of Galactic molecular clouds. Table 5 lists the observed flux densities for DR 21(OH), DR 21 West, and W33 Met in various class I lines, as well as the flux densities for models selected as the best fits according to a χ^2 criterion[3]. The calculated models describe the observed line intensity ratios for E-methanol fairly well. The models for the $7_0{-}6_1A^+$ line underestimate the flux densities compared to the observed values. This may be due to the simplicity of the models, in which we assumed that different lines are emitted in the same region. In reality, the masing region may have multiple components (or multiple layers), and the emission in different lines may form in different layers with different physical conditions.

5.3. Class II Masers

In these observations, we detected class II masers only in the $0_0{-}1_{-1}E$ line toward G345.01+1.79 and possibly M8E. We found only thermal sources In the $3_{-2}{-}4_{-1}E$ and $J_1{-}J_0E$ lines.

As noted above, class II masers are generated under the action of strong external radiation. This can be emission from hot dust or free–free radio emission from ultracompact HII regions. We have modeled the excitation of methanol by radio emission from HII regions. As an example, Table 6 lists the results of LVG modeling of a class II maser source. In this model, the gas density was 3×10^6 cm^{-3}, temperature 50 K, and methanol density divided by the velocity gradient 0.5×10^{-1} cm^{-3}/(km/s pc^{-1}). The source is observed against the background of a compact HII region with an emission measure of 3×10^{12} cm^{-6} pc. The dilution factor for the HII region was taken to be 0.5. For comparison, we list flux densities in maser lines of the class II source G345.01+1.79; the model flux density in the $5_1{-}6_0A^+$ line was taken to equal the observed value, and the flux densities in other lines were calculated from the ratios of the model brightness temperatures in a way similar to that used when modeling the class I masers. A number of models taking into account emission by hot dust were calculated by Sobolev *et al.* [6]; the results for their model no. 6 are also given in Table 6.

The models for pumping by free–free emission from a compact HII region and by emission from hot dust are both in reasonable agreement with the observational data. The agreement with the flux at 6.7 GHz was predetermined; the predicted fluxes in the lines at 12.2 GHz, 107.0 GHz, 157.2 GHz agree with the observations fairly well. The models overestimate the fluxes for the lines at 108.9 GHz and 156.6 GHz. The largest discrepancy with the observations is shown by the predictions of the HII

---

[3] The LVG method enables calculation of the line brightness temperature, but does not yield source sizes, so that we could not use it to estimate the flux densities. Therefore, we ascribed the observed value to the model flux density in the $4_{-1}{-}3_0E$ line and determined the flux densities in the remaining lines from the ratios of the model brightness temperatures, assuming the sizes of each source in all the lines to be identical.

model for the lines at 165.2 and 230.0 GHz: the model predicts large fluxes, which are certainly not observed. At the same time, the dust model predicts an absence of an appreciable flux at these frequencies, in agreement with the observations. Thus, the model for pumping by emission from hot dust describes the observations better. However, the difference between the HII and dust models is purely formal, since real HII regions also contain hot dust mixed with the ionized gas.

Thus, the results of our observations lead us to conclude that modeling of class II methanol masers should take into account pumping by emission from hot dust.

6. CONCLUSIONS

As a result of our 1-mm, 2-mm, and 3-mm observations, we have detected both maser and thermal methanol sources. Methanol masers have been detected at 1 mm for the first time. We found a unique maser in the 0_0–$1_{-1}E$ line at 3 mm.

Masers in the 8_{-1}–7_0E line that are comparable in intensity with masers in the lines 4_{-1}–3_0E, 5_{-1}–4_0E and other lines belonging to class I can arise at gas temperatures of 50–60 K and densities of 3–6×10^4 cm^{-3}, typical of Galactic molecular clouds.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Gabriel Paubert and other colleagues from the Pico Veleta Observatory for help with the observations. This work was partially supported by the Soros Foundation (grant MND300) and the Russian Foundation for Basic Research (project no. 95-02-05826).

REFERENCES

1. J. A. Ball, C. A. Gottlieb, A. E. Lilley, and H. E. Radford, Astrophys. J. Lett. **162**, L203 (1970).
2. A. J. Beasley, S. P. Ellingsen, M. J. Claussen, and E. Wilcots, Astrophys. J. **459**, 600 (1996).
3. W. Batrla, H. E. Matthews, K. M. Menten, and C. M. Walmsley, Nature **326**, 49 (1987).
4. K. M. Menten, in *Proceedings of the 3rd Haystack Observatory Meeting "Skylines,"* Ed. by A. D. Haschick and P. T. P. Ho, Astron. Soc. Pac. Conf. Ser. **16**, 119 (1991).
5. D. M. Cragg, K. P. Johns, P. D. Godfrey, and R. D. Brown, Mon. Not. R. Astron. Soc. **259**, 203 (1992).
6. A. M. Sobolev, D. M. Cragg, and P. D. Godfrey, Mon. Not. R. Astron. Soc. **288**, L39 (1997).
7. M. A. Voronkov, Pis'ma Astron. Zh. **25**, 86 (1999) [Astron. Lett. **25**, 149 (1999)].
8. D. M. Cragg, M. A. Mekhtiev, and R. P. A. Bettens, Mon. Not. R. Astron. Soc. **264**, 769 (1993).
9. I. E. Val'tts, Astron. Zh. **76**, 189 (1999) [Astron. Rep. **43**, 157 (1999)].
10. I. E. Val'tts, Pis'ma Astron. Zh. **24**, 910 (1998) [Astron. Lett. **24**, 788 (1998)].
11. J. L. Caswell, R. A. Vaile, S. P. Ellingsen, and R. P. Norris, Mon. Not. R. Astron. Soc. **274**, 1126 (1995).
12. J. L. Caswell, Mon. Not. R. Astron. Soc. **274**, 215 (1998).
13. I. E. Val'tts, S. P. Ellingsen, V. I. Slysh, *et al.*, Mon. Not. R. Astron. Soc. **310**, 1077 (1999).
14. V. I. Slysh, S. V. Kalenskii, and I. E. Val'tts, Astrophys. J. **442**, 668 (1995).
15. V. I. Slysh, S. V. Kalenskii, I. E. Val'tts, and V. V. Golubev, Astrophys. J. Lett. **478**, L37 (1997).
16. L. R. Kogan and V. I. Slysh, Astrophys. J. **497**, 800 (1998).
17. A. D. Haschick, K. M. Menten, and W. Baan, Astrophys. J. **354**, 556 (1990).
18. V. I. Slysh, S. V. Kalenskiĭ, and I. E. Val'tts (2001) (in press).
19. V. I. Slysh, S. V. Kalenskii, I. E. Val'tts, *et al.*, Astrophys. J., Suppl. Ser. **123**, 515 (1999).
20. S. Liechti and T. L. Wilson, Astron. Astrophys. **314**, 615 (1996).
21. S. V. Kalenskiĭ, V. I. Slysh, I. E. Val'tts, *et al.*, Astron. Zh. **78**, 31 (2001) [Astron. Rep. **45**, 26 (2001)].
22. V. I. Slysh, S. V. Kalenskii, and I. E. Val'tts, Astrophys. J. Lett. **413**, L133 (1993).
23. I. E. Val'tts, A. M. Dzyura, S. V. Kalenskiĭ, *et al.*, Astron. Zh. **72**, 22 (1995) [Astron. Rep. **39**, 18 (1995)].
24. R. L. Plambeck and K. M. Menten, Astrophys. J. **364**, 555 (1990).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 57–66. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 63–72.
Original Russian Text Copyright © 2002 by Lekht, Mendoza-Torres, Berulis.

Cyclic Activity of the Water-Vapor Maser in S128

E. E. Lekht[1,2], **J. E. Mendoza-Torres**[1], **and I. I. Berulis**[3,4]

[1] *Instituto Nacional de Astrofísica, Óptica y Electrónica,*
Luis Enrique Erro No. 1, Apdo Postal 51 y 216, Tonantzintla, Puebla, 72840 México
[2] *Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899 Russia*
[3] *Pushchino Radio Astronomy Observatory, Astro Space Center,*
Lebedev Institute of Physics, Russian Academy of Sciences,
Pushchino, Moscow oblast, 142292 Russia
[4] *Kaunas University of Technology, K. Donelaičio g. 73, Kaunas, 3006 Lithuania*

Received April 22, 2001

Abstract—The results of observations of the S128 H_2O maser carried out from February 1995 to March 2001 on the 22-m radio telescope of the Pushchino Radio Astronomy Observatory are presented. Two activity cycles of the H_2O maser with a period of about 10 years were observed during the total monitoring interval (1981–2001). This may be connected either with cyclic activity of the central star in S128 during its formation or with the influence on the H_2O masering region of shocks arising near an ionization front at the interface of two colliding CO clouds. The emission at radial velocities from −73 to −70 km/s consists of four emission features. The emission feature at −71.8 km/s exhibits a flux dependence on linewidth that is typical of an unsaturated maser. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The existence of cyclic variability of water-vapor masers in star-forming regions was first demonstrated by Lekht and Sorochenko [1] and Liljeström *et al.* [2]. Subsequent long-term monitoring yielded periods for the long-term variations of the total flux of H_2O masers that were typically 6–18 years (see, for example, [3, 4]).

The source S128 holds a special position among H_2O masers associated with star-forming regions. Its distinguishing feature is that it lies at an interface between two clouds that are probably colliding. The S128 region has been mapped in the CO line [5]. The boundary separating the clouds lies in the north–south direction. To the west of this boundary is a cloud with velocity −71.5 km/s, which has a small radial-velocity gradient; to the east is a cloud with −74 km/s, which, on the contrary, has a large velocity gradient near the cloud interface. The compact HII region S128N ($\sim 3''$) and, $60''$ to the south of it, the extended HII region S128 [6] are located along this boundary.

The H_2O maser is located $12''$ to the north of S128N [6]. Later, it was shown that there are two water masers separated by a projected distance of $\sim 13''$ [5]. VLBI observations [7] carried out in June, 1996, confirmed the existence of two maser sources, separated by $\approx 13''$. The accepted kinematic distance to S128 is 7.5 kpc [6]. To illustrate the structure of the S128 region, we present a schematic map in Fig. 1. The data for this figure were taken from Haschick *et al.* [5]. We can see that both HII regions and the masers $H_2O(A)$ and $H_2O(B)$ are all located along the boundary between the two colliding CO molecular clouds; i.e., near an ionization front (dashed curve). The positions of the infrared sources IRS1 and IRS2 are also plotted [8]. The cloud collision could initiate star formation in S128, and specific conditions stimulating star formation could be present here. The preferred mechanism is triggered star formation.

The present work continues studies of the evolution of the H_2O maser emission in S128 on the basis of monitoring data obtained in 1995–2001. In addition, we use the data of [9] for 1981–1994.

2. OBSERVATIONS AND DATA

Observations of the H_2O maser emission toward S128 ($\alpha_{1950} = 21^h30^m37^s$, $\delta_{1950} = 55^\circ40'36''$) were carried out on the 22-m radio telescope of the Pushchino Radio Astronomy Observatory between February 1995 and March 2001. The antenna beamwidth is $2.6'$, so we received emission from both maser sources in S128. The average interval between observations was 1.5 months. The system had a cooled transistor amplifier at its front end and a noise temperature of 150–230 K. In September, 2000, the

1063-7729/02/4601-0057$22.00 © 2002 MAIK "Nauka/Interperiodica"

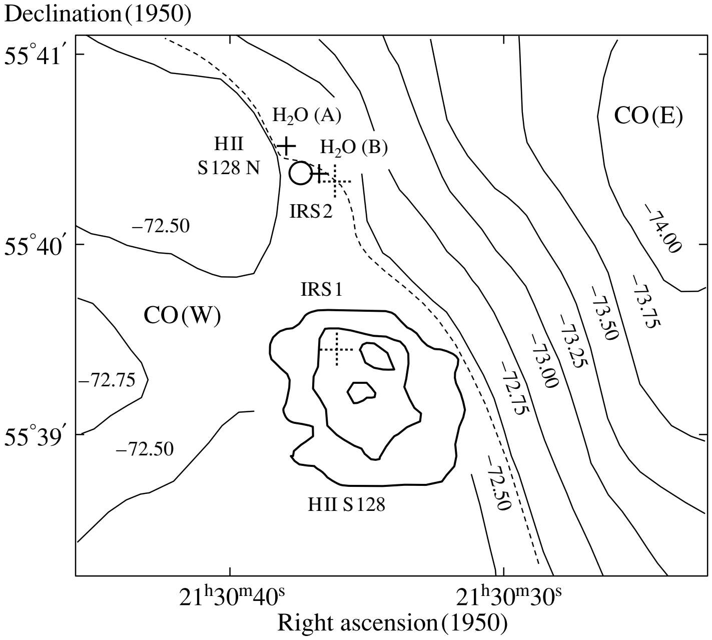

Fig. 1. Schematic map of the S128 region showing contours of equal ^{12}CO radial velocities with the velocity for each contour in km/s marked, HII regions, H_2O masers, and IR emission sources. The dashed curve shows the probable boundary separating the CO clouds. The solid and dotted crosses show the positions of H_2O masers and infrared sources (IRS), respectively. The sizes of the dotted crosses show the position errors.

radiometer was upgraded, making it possible to lower the system noise temperature to 100–130 K. The signal analysis was carried out using a 96-channel (before July 1997) or 128-channel (starting in July 1997) filter-bank spectrum analyzer with a resolution of 7.5 kHz (0.101 km/s in radial velocity at 22 GHz).

The results of our 1995–2001 observations of the H_2O maser S128 are presented in Fig. 2. The horizontal axis plots the radial velocity in km/s with respect to the Local Standard of Rest and the vertical axis the flux density in Janskys (the scale in Janskys is shown by the vertical arrow). For convenience, we present all spectra on the same scale. Figure 2g shows a superposition of all our spectra obtained in this interval.

Figure 3 presents in three dimensions the evolution of the S128 maser emission throughout the monitoring interval (1981–2001). The x, y, and z axes plot time in years, radial velocity in km/s, and flux density in Janskys, respectively. The top part of the figure gives contours of equal flux in the time–velocity plane. The contours are drawn in intervals of 10%, and also at 2 and 5% of the maximum flux. The peaks of burstlike emission are somewhat lower than the peaks in the spectra. This is a result of the specific 3D-imaging method we have used, which requires a uniform data series along the x and y axes. Since this condition is not fulfilled on the time axis, we have interpolated the spectra so that they are equally spaced in time. These interpolated spectra were then used to construct the three-dimensional image. In one period of almost two years, from May, 1993, to January, 1995, there was only one observation (March, 1994). For this reason, the interpolation in this interval led to the appearance of a small "tail" on the feature at—80 km/s.

Figure 4 shows variations of the total (integrated) H_2O line flux. We have also used data for 1979–1980 [6] and 1981–1994 [9] when plotting this figure. Two main cycles of the total H_2O line flux variations are visible. We have fitted a smoothed curve (dashed) to each cycle. The variability period is about 10 years. In addition, there are more rapid variations of the total flux with a timescale of two to four years. Due to these variations, each activity cycle has three maxima: preliminary (preceeding the main maximum), main, and secondary. The time intervals between these maxima are the same in both cycles.

During the entire 20 years of our monitoring of S128, we observed only three short flares in the emission of individual features. The first flare occured at the end of 1981 and beginning of 1982 ($V_{\rm LSR} = -73.5$ km/s, $F = 460$ Jy). The strongest flare was observed in the second half of 1985, at $V_{\rm LSR} = -77.3$ km/s (Fig. 3); its peak emission flux reached 1730 Jy. The estimated lifetime of this component was ~ 6 months. The third flare took place in the

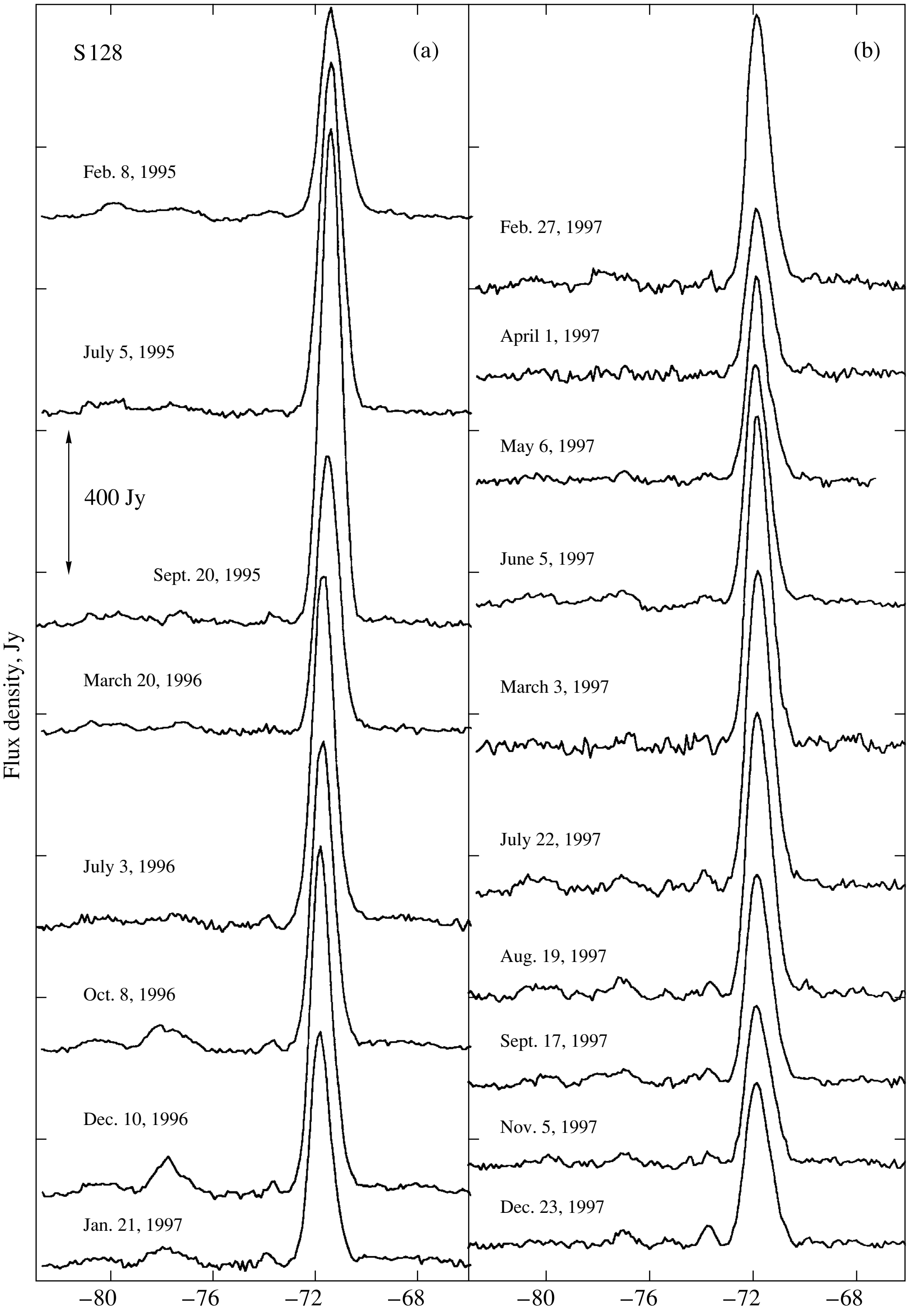

Fig. 2. (a)–(f) Spectra of the H_2O maser emission of S128 observed in 1995–2001 and (g) a superposition of all spectra for this time interval.

middle of 1986 at $V_{\rm LSR} = -78$ km/s and had a flux of 750 Jy.

The average spectrum for 1995–2001 is shown in Fig. 5. For comparison, we also present the average spectrum for 1981–1994 [9] (dashed) and the difference between these two spectra (dotted). The

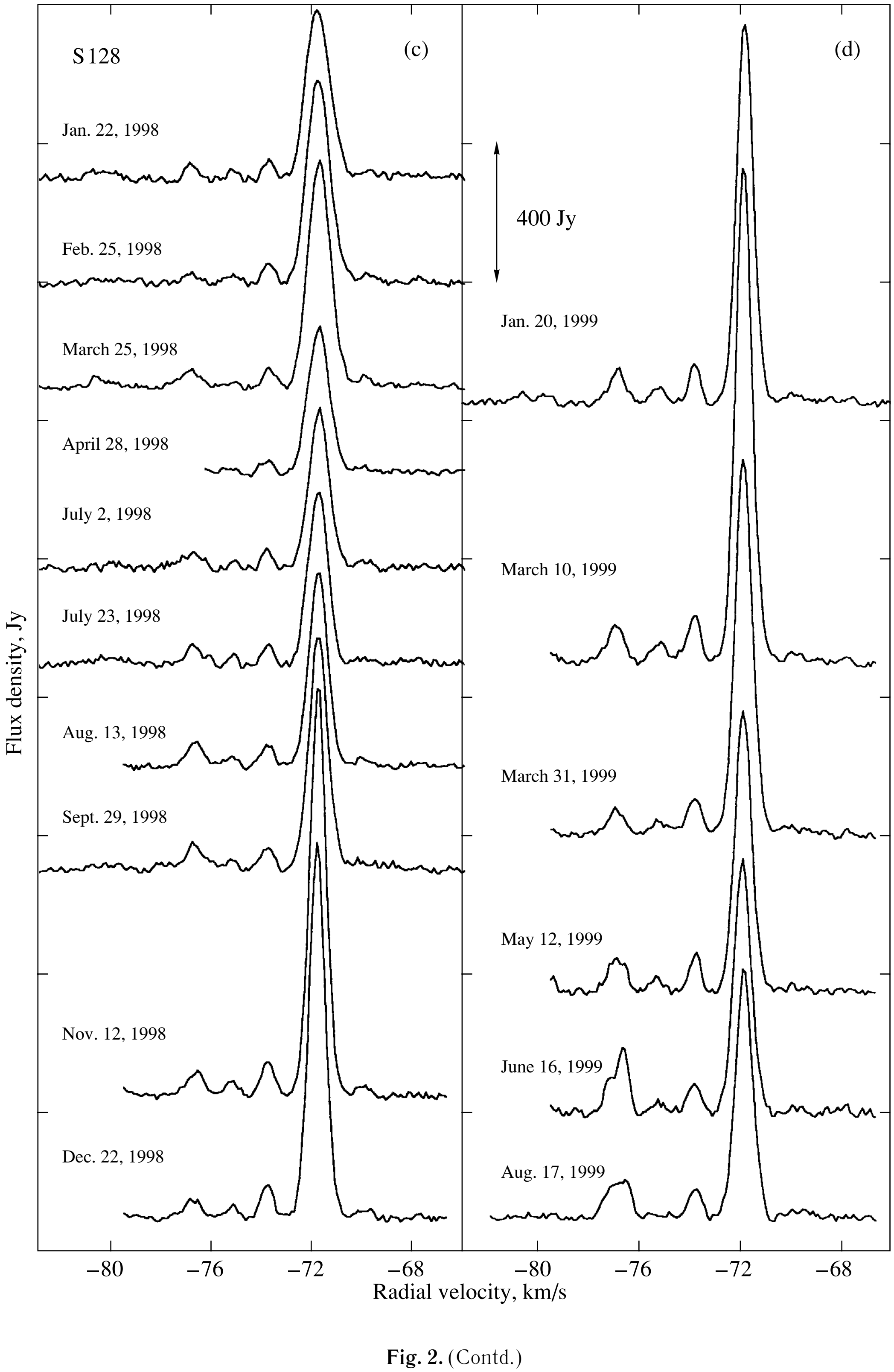

Fig. 2. (Contd.)

structure of the average spectra for the two intervals is fairly similar. We can clearly see that, as earlier [9], the emission features form three main groups in the radial-velocity intervals −81 to −76, −76 to −73, and −73 to −68 km/s. However, the emission has been redistributed among the groups of features

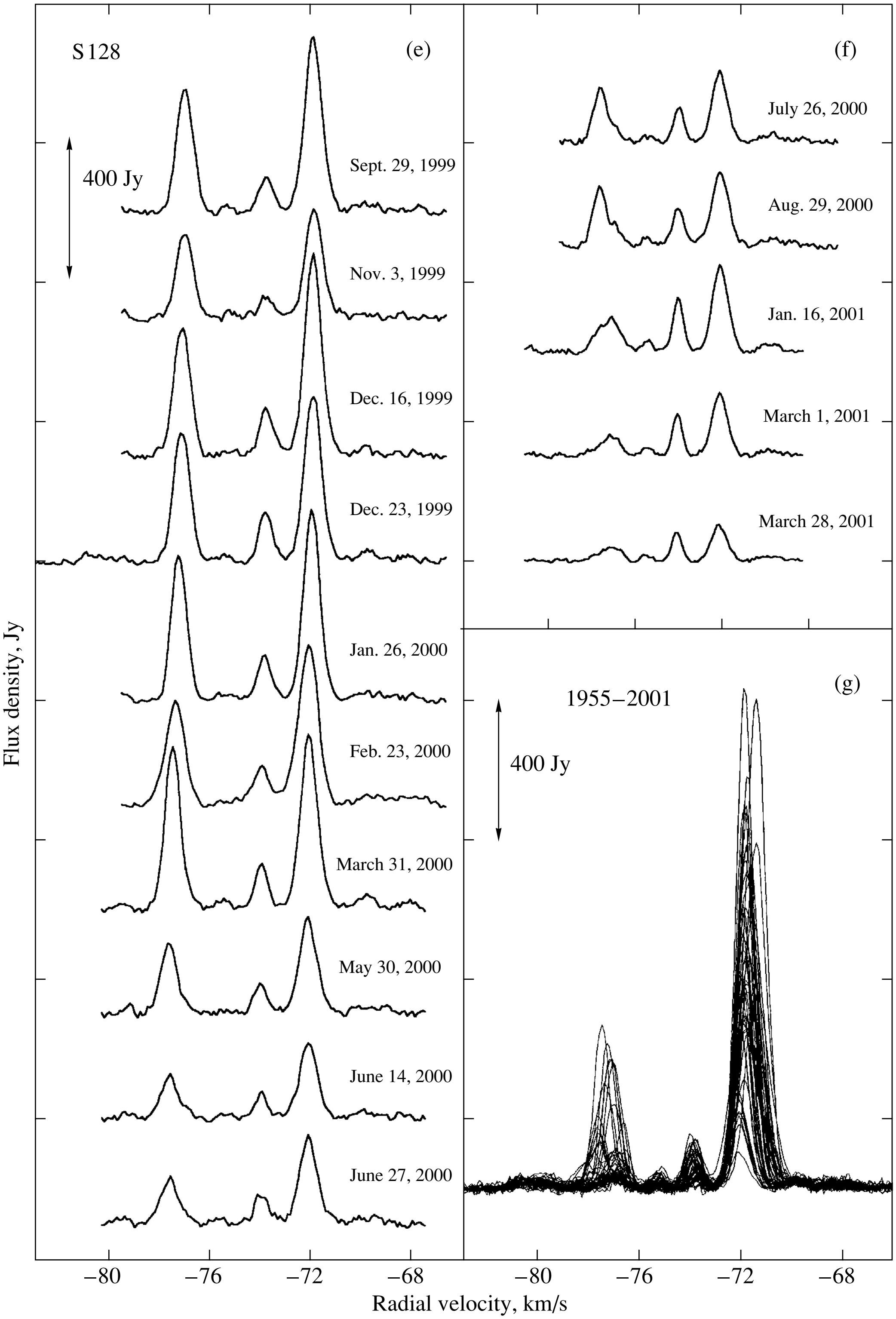

Fig. 2. (Contd.)

(dashed curve in Fig. 5). The total flux for the 1981–1994 average spectrum is 700 Jy km/s, while that for the 1995–2001 spectrum is 860 Jy km/s; i.e., the difference is small. From time to time, a triplet structure was observed, for example, in 1983–1984 and 1999–2000. All these properties are reminiscent

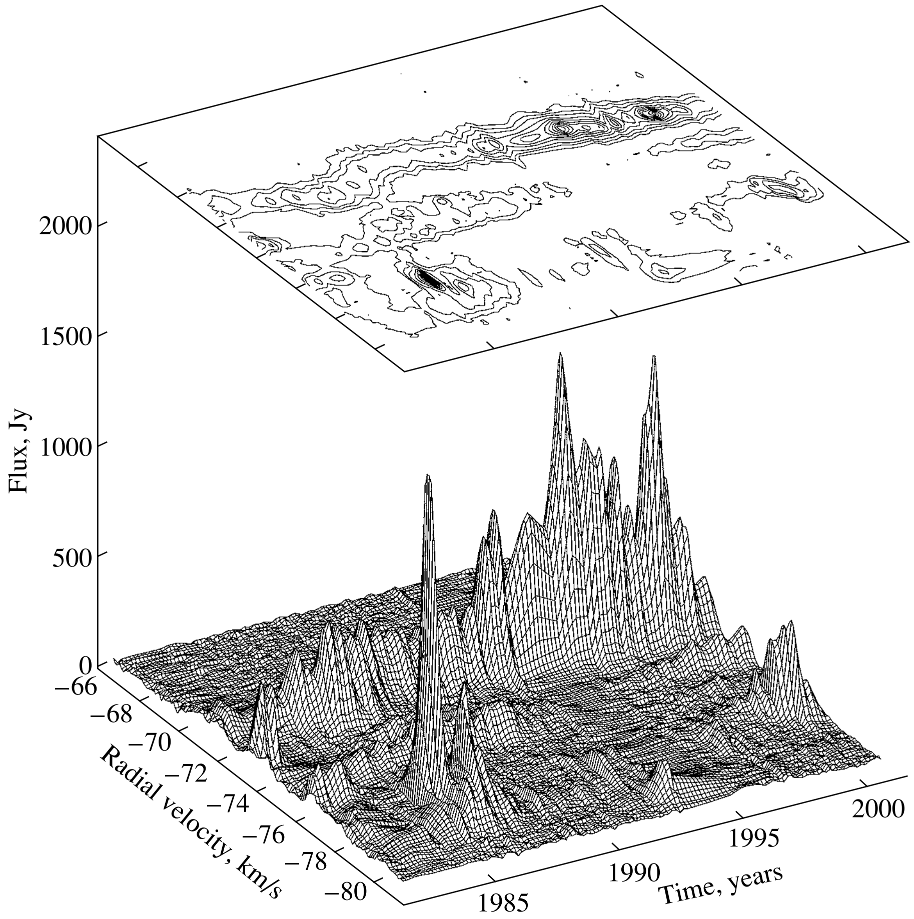

Fig. 3. Three-dimensional representation of the evolution of the H_2O maser emission in S128 (see text).

of the source S269, which is probably identified with a Keplerian disk [10].

Note that the emission feature at $V_{LSR} \approx -71.5$ km/s was always present. It was usually the most intense feature and, as a rule, represented a superposition of two components with close radial velocities. Figure 6 shows the variations of the flux, peak radial velocity, and linewidth of this feature in 1987–2001. The vertical dashed lines denote boundaries separating this into subintervals. This division was primarily based on the flux variations, but we also took into account the radial-velocity and linewidth variations. We divided the total interval into seven parts, which we have labelled a to g. Between d and e, there is a nearly empty span of two years, with only one observation in March, 1994. The smoothed curve in the top part of Fig. 6 (dotted) was fitted in order to delineate long-term and rapid (burstlike) flux variations. The arrows mark the main flare components.

We found a correlation between the flux and linewidth variations for the feature that had a radial velocity of −71.8 km/s between April, 1998, and September, 1999. The open and filled circles in Fig. 7 correspond to points of the ascending and descending branches of the flux curve, respectively. We fit smoothed curves for both branches, shown by the dashed and dot–dashed curves. The dotted curve shows the smoothed curve derived for the entire data set.

3. DISCUSSION

We have used almost 20 years of monitoring data for S128 in the 1.35-cm water-vapor line to study variations of the total flux, as well as the evolution of the main emission feature (with an average radial velocity of about −71.5 km/s) in 1987–2001. We have obtained curves describing the flux, radial-velocity, and linewidth variations of this spectral component. In addition, we searched for correlations between the variations of these parameters.

3.1. Variability of the Integrated Flux

Our long-term monitoring of the S128 H_2O maser has revealed important regularities in the variations of

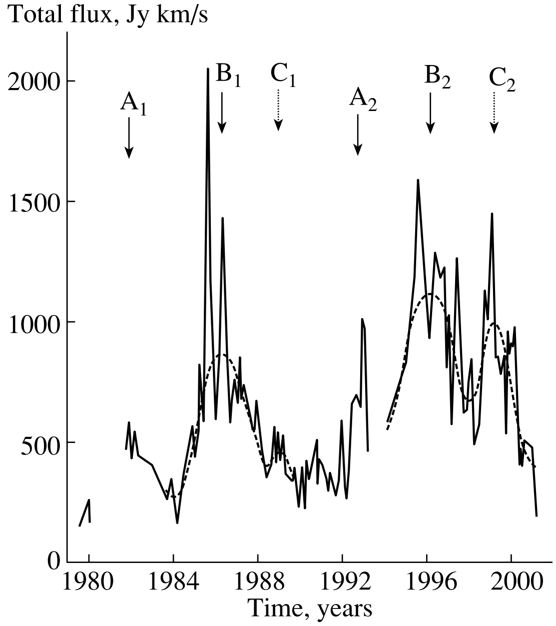

Fig. 4. Variations of the integrated flux of the H_2O maser emission in S128. The arrows with labels mark maxima of the total flux found by fitting smoothed curves (dashed) to the data.

the total maser flux. There were two maser activity cycles with intervals between their maxima of about 10 years (Fig. 4). We also observed more rapid variations of the total flux with time scales of two to four years (dashed curves in Fig. 4).

In spite of the fact that the emission of the third group of features (−73 to −69 km/s) dominated in the second activity cycle, the character of the total-flux variability remained the same: the three maxima are arranged in the same way as in the first activity cycle.

The long-term variability of the maser emission with a period of about ten years may be associated with the non-stationary character of the formation of the central star in S128. However the situation is much more complicated here than in other masers for two main reasons. First, we receive the flux from both H_2O masers in S128, whose angular separation is 13″. For a distance to S128 of 7.5 kpc, this corresponds to about 3 pc. Second, and also very important, both masers and the HII regions are located at an interface of two interacting clouds, near an ionization front. Note also that both H_2O masers and the HII region S128N lie at the position of the ^{12}CO emission peak.

The projected distance of 3 pc between the maser sources is too large for a central star to be a common activity center for both. Therefore, we cannot exclude the possibility that each of the H_2O masers has its own activity center. A common origin of the maser variations in both sources could be shocks arising near the ionization front. Maxima B and C were more pronounced in the second activity cycle, when the emission at radial velocities of −73 to −70 km/s dominated. According to Haschick *et al.* [6] and Migenes *et al.* [7], this emission comes mostly from the maser H_2O(B). This suggests that the variability on time scales of two to four years is associated with H_2O(B).

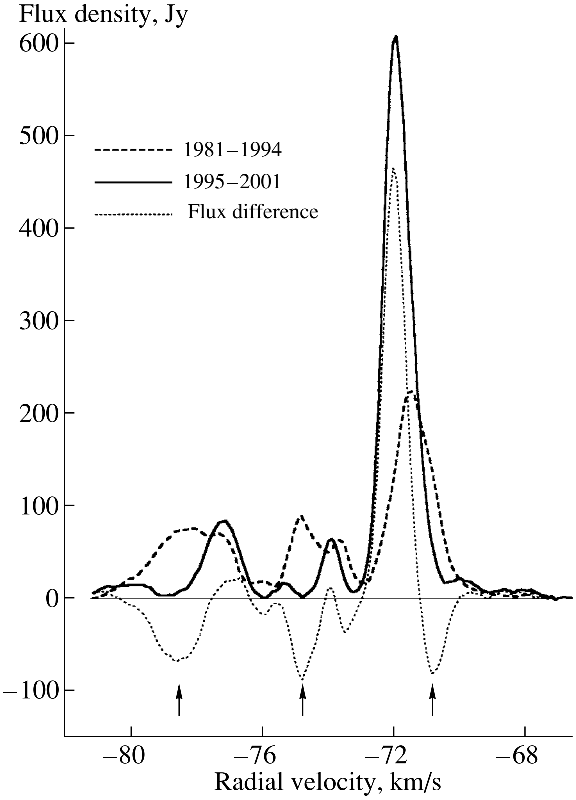

Fig. 5. Average spectra of the S128 H_2O maser for two activity cycles: 1981–1994 (dashed) and 1995–2001 (solid). The dotted curve shows the difference between these two spectra. The arrows at the bottom mark features observed only in the first maser actvity cycle.

3.2. Average H_2O Spectra

Early in our monitoring, the H_2O spectra of S128 had a more-or-less triplet structure. Beginning in 1988, the spectrum consisted mainly of the central (−76 to −71 km/s) and right (−71 to −69 km/s) groups of features. Later, the emission of the right group dominated. Since 1999, the spectrum has again acquired a triplet structure. Beginning in the middle of 2000, we observed a rapid decline of the intensities of all spectral components.

To study the evolution of the H_2O spectrum, we obtained average spectra for the two maser activity cycles of S128, for 1981–1994 and 1995–2001

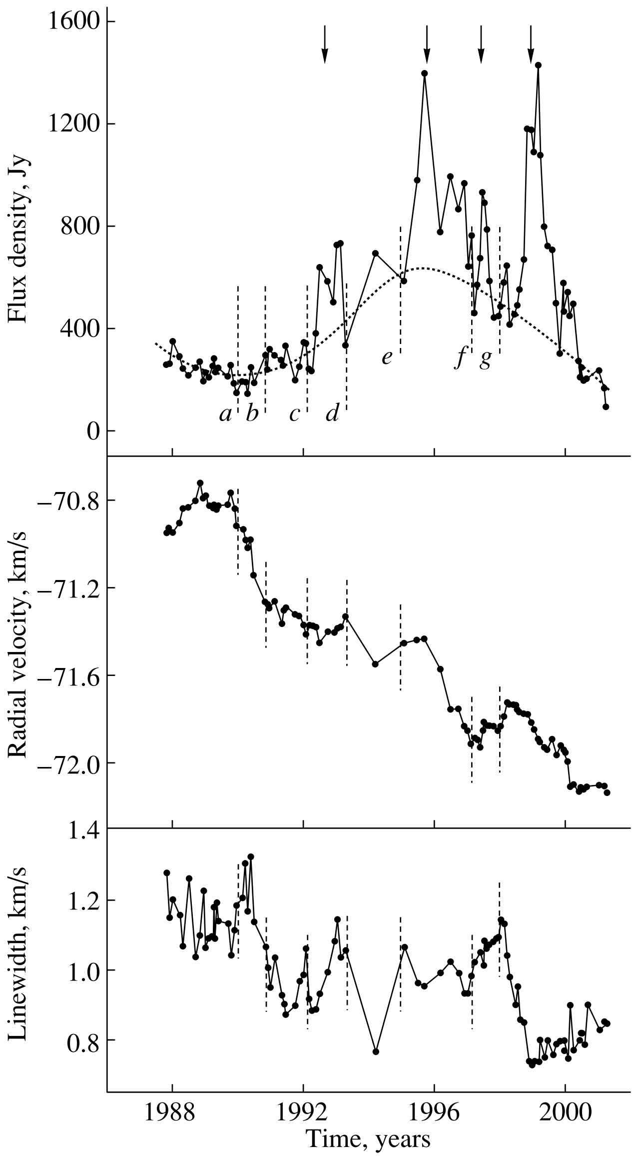

Fig. 6. Variations of the flux, radial velocity, and linewidth of the −71.5-km/s feature of the S128 H_2O maser in 1987–2001. The fitted curve (dotted) describes the long-term flux variations. The arrows mark the main intervals with rapid flux variations. The vertical dashed lines divide the entire time interval into subintervals, labeled with the letters shown.

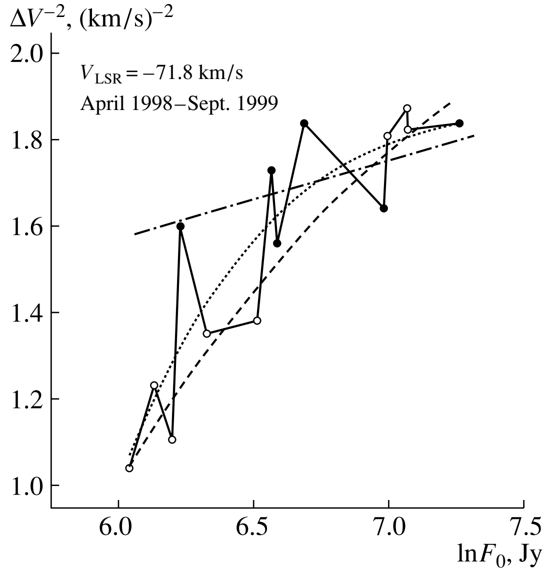

Fig. 7. The flux F_0 as a function of the linewidth ΔV of the main feature of the S128 H_2O maser for April, 1998–September, 1999. The open and filled circles correspond to the ascending and descending branches of the flux curve, respectively. The point for the flux maximum belongs to both branches. The dashed, dot–dashed, and dotted curves show the fitted curves for the ascending branch, descending branch, and the entire set of points, respectively.

(Fig. 6). The structures of these spectra are both similar and difference. In both cases, there is the same division into three groups of features. Furthermore, the total radial-velocity range remained unchanged, and the total fluxes have changed only slightly, being 700 and 860 Jy km/s for the first and second maser activity cycles, respectively.

The main difference between the two average spectra is that the emission of the components at −78.5, −74.8, and −70.8 km/s has vanished in the later spectrum. Given the small change in the total flux, this means that the emission was redistributed among the components: the emission of the third group of features, at the higher radial velocities, has become more intense.

The structure of the average spectra at −73 to −70 km/s shows that there are three main emission features with radial velocities of −71.8, −71.4, and −70.8 km/s.

3.3. Maser Emission at −73 to −70 km/s

We divided the 1987–2001 data into subintervals in order to study the spectrum's structure and search for correlations between variations of the flux, radial velocity, and linewidth. In the interval a–d, we observed no significant changes in the flux, and note only the flare of 1992. In this case, we observed a broadening of the −71.4-km/s feature from approximately 0.9 to 1.1 km/s. Since the line was asymmetric and fairly broad, this was probably due to the presence of a blending feature to the right.

The curve describing the emission-peak radial-velocity drift in Fig. 6 has four plateaus. The transition from one plateau to another occurs rather smoothly, without sharp jumps. This property of the variability in the velocity interval from −73 to

−70 km/s is due to the existence of four emission features, at −72.1, −71.8, −71.4, and −70.8 km/s. A consecutive shift of the emission peak from lower to higher radial velocities indicates an ordered arrangement of emitting structures in the source.

The results obtained coincide with those of Section 3.2. Only one new feature appeared, at −72.1 km/s; it was dominant only starting in 2000, during the overall decline of the maser activity, so that it did not make an appreciable contribution to the average spectrum. There is good agreement with the data of Comoretto *et al.* [11] and Migenes *et al.* [7]. The radial velocity of this feature was −70.8 km/s in July, 1988 [11], and −71.7 km/s in June, 1996 [7].

Thus, there is evidence that the medium responsible for the maser emission at radial velocities from −73 to −70 km/s is strongly fragmented. The fragments probably form an ordered structure (possibly part of a shell), with the maser emission induced by a common source—the central star. The central velocity of this interval coincides with that of the western CO cloud, −71.5 km/s (Fig. 1). It is quite possible that the spectrally compact group of features forming the maser H_2O(A) belongs to this cloud, in which there is no large radial velocity gradient [5].

3.4. The Emission Feature at −71.8 km/s

The H_2O line was symmetric for two features of the third group, at −71.8 and −72.1 km/s. The former feature was observed at the maximum of activity of the S128 maser, and the latter, during the activity decline. Our analysis of the variations of the radial velocity, linewidth (Fig. 6), and line shape has shown that the observed variability of these parameters cannot be a consequence of changes in the flux ratios for the −71.8 and −72.1 km/s features. We observed essentially a single feature at −71.8 km/s from April, 1998, to September, 1999. Its small velocity drift (about 0.2 km/s) may be real. There was a strong flare of the emission in this period. In addition to the flux variations, there were considerable changes in the linewidth that were correlated with the flux variations. The line became narrower on the ascending branch of the flux curve and broader on the descending branch. These changes can be described by the dependence

$$\Delta V^{-2} \propto \ln F_0, \qquad (1)$$

where ΔV is the line full width at half maximum and F_0 is the flux at the line center.

There are two branches of this dependence: one for the flux growth and the other for its decline. The line broadening on the descending branch proceeded slower than the narrowing during the ascending branch (Fig. 7). In addition, the flux dependence of the linewidth for the ascending branch differs slightly from (1) (dashed curve in Fig. 7). It can readily be shown that, in this case

$$\Delta V^{-2.2} \propto \ln F_0. \qquad (2)$$

This relationship between the flux and linewidth indicates that the S128 H_2O maser associated with the emission feature at −71.8 km/s is unsaturated [12].

4. CONCLUSION

Our analysis of the S128 H_2O maser spectra using data obtained in 1982–2001 has yielded the following results.

(1) In 1982–2001, the water-vapor maser in S128 went through two activity cycles, with 10 years between the flux maxima. We observed three flux maxima in the same time sequence in each cycle: preliminary, main, and secondary.

(2) The emission features in the H_2O spectrum form three groups, with the boundaries between them at −76 and −73 km/s.

(3) The structure of the average spectra in the two variability cycles was similar, indicating the presence of fairly stable structures in S128. In some time intervals, the spectra had a triplet structure (e.g., in 1999–2000).

(4) The ranges of the maser-emission radial velocities for two time intervals (1981–1993 and 1995–2001) were the same, −81 to −67 km/s.

(5) In 1987–2001, the peak of the component at $V_{\rm LSR} \approx -71.5$ km/s drifted in radial velocity from −70.8 to −72.1 km/s. The origin of this drift was the consecutive appearance of four emission features at −70.8, −71.4, −71.8, and −72.1 km/s.

(6) In 1998–1999, we observed a linewidth–flux dependence typical of an unsaturated maser for the emission feature at −71.8 km/s. The line shape remained symmetrical.

ACKNOWLEDGMENTS

The Pushchino Radio Astronomy Observatory RT-22 radio telescope is supported by the Ministry of Industry, Science, and Technology of the Russian Federation (formerly the State Committee for Science and Technology of the Russian Federation; registration number 01-10). This work was supported by the Russian Foundation for Basic Research (project no. 99-02-16293). The authors are grateful to the staff of the Pushchino Radio Astronomy Observatory for help with the observations.

REFERENCES

1. E. E. Lekht and R. L. Sorochenko, in *Proceedings of the XIX All-Union Conference "Galactic and Extragalactic Radio Astronomy," [in Russian] Tallinn, 1987,* p. 118.
2. T. Lijeström, K. Mattila, M. Toriseva, *et al.*, Astron. Astrophys., Suppl. Ser. **79**, 19 (1989).
3. E. E. Lekht, Astron. Zh. **73**, 695 (1996) [Astron. Rep. **40**, 632 (1996)].
4. E. E. Lekht, Astron. Astrophys., Suppl. Ser. **141**, 185 (2000).
5. A. D. Haschick and P. T. P. Ho, Astrophys. J. **292**, 200 (1985).
6. P. T. P. Ho, A. D. Haschick, and F. P. Israel, Astrophys. J. **243**, 526 (1981).
7. V. Migenes, S. Horiuchi, V. I. Slysh, *et al.*, Astrophys. J. Suppl. Ser. **123**, 487 (1999).
8. A. Mampaso, P. Gómez, C. Sánchez-Magro, *et al.*, Mon. Not. R. Astron. Soc. **207**, 465 (1984).
9. I. I. Berulis, E. E. Lekht, and E. Mendoza-Torres, Astron. Zh. **72**, 468 (1995) [Astron. Rep. **39**, 411 (1995)].
10. E. E. Lekht, M. I. Pashchenko, and I. I. Berulis, Astron. Zh. **78**, 1081 (2001) [Astron. Rep. **45**, 949 (2001)].
11. G. Comoretto, F. Palagi, R. Cezaroni, *et al.*, Astron. Astrophys., Suppl. Ser. **84**, 179 (1990).
12. K. Matilla, N. Holsti, M. Toriseva, *et al.*, Astron. Astrophys. **145**, 192 (1985).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 67–73. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 73–80.
Original Russian Text Copyright © 2002 by A. Podgorny, I. Podgorny.

Numerical MHD Simulations of Post-Flare Loop Formation on the Sun: Allowing for Thermal-Conductivity Anisotropy

A. I. Podgorny[1] **and I. M. Podgorny**[2]

[1]*Lebedev Institute of Physics,*
Russian Academy of Sciences,
Leninskiĭ pr. 53, Moscow, 117924 Russia
[2]*Institute of Astronomy,*
Russian Academy of Sciences,
Pyatnitskaya ul. 48, Moscow, 109017 Russia

Received February 7, 2001

Abstract—The process of post-flare loop formation, including the heating of flux tubes by hot chromospheric sources and their filling with plasma, is demonstrated by simulations in an MHD approximation. The loop is additionally heated at its apex by the interaction of oppositely directed plasma streams. Local coronal heating over the loop is also possible due to magnetic-field-line reconnection. A new version of the PERESVET code that can take into account anisotropy of the thermal conductivity of a plasma in a magnetic field was used for the computations. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The post-flare formation of a loop is among the most impressive events accompanying a solar flare. Bright loops aligned with magnetic-field lines can reach heights of $\sim 10^{10}$ cm. Loops can be observed in various spectral ranges. They are especially well-defined in the emission lines of multi-charge ions (OV, MgIV, FeXVI, etc.). In many cases, the loop thickness does not exceed 5×10^8 cm. The plasma densities in a post-flare loop are two to three orders of magnitude higher than the coronal density, and the temperatures reach 700 eV. The outer part of a loop is typically hotter than its inner part [1]. The loops are anchored in the chromosphere, which is most likely the source of the dense plasma in the hot loops. The possible ejection of highly ionized material from the feet of a loop was demonstrated in [2–4]. Culhane *et al.* [4] propose that such impulsive chromospheric evaporation is due to the release of energy by fast electrons precipitating into the chromosphere during a flare. In the electrodynamic flare model [5], fast electrons are accelerated in an upward field-aligned current. As this takes place, two centers of chromospheric evaporation form at the two feet of the loop, and the plasma streams flowing from these centers fill the magnetic flux tube. This pattern of post-flare loop formation has been confirmed in numerical MHD simulations [6], which have shown that the formation of the chromospheric sources results in rapid filling of the flux tube with dense plasma. Streams from the two sources interact at the loop apex. The width of the interaction region may be determined by the mean free path of ions or by various kinetic processes that are neglected in an MHD approximation. "Cool" loops are observed along with hot ones. They can clearly be seen in the H_α neutral-hydrogen line. There is probably condensation of material in these cool loops; we do not consider this process here.

As very powerful chromospheric sources arise, the magnetic field at the apex of the loop may cease to confine the plasma streams; in this case, the magnetic-field lines will stretch upward to form a current-sheet configuration, and the merging streams from the two sources will be ejected from the corona into interplanetary space in the form of a transient. Such processes have been observed on the Sun [7] and simulated numerically [8].

The main drawback of previously published numerical MHD simulations of coronal events is their neglect of anisotropy of the plasma's thermal conductivity in a magnetic field. The choice of a Spitzer thermal conductivity ($\sim T^{5/2}$) in the heat-conduction equation implies rapid cooling of the plasma streams in the loop (due to heat losses to the ambient coronal plasma) and rapid heating of the cool loops. Therefore, such numerical computations can demonstrate the filling of a loop with plasma, but cannot adequately simulate thermal effects. In analytical calculations, this difficulty is obviated by assuming the plasma to be thermally conductive only in the direction of the magnetic field, which reduces the problem to a one-dimensional form. However, the one-dimensional approximation fails to take into account plasma-dynamical effects of field-line curvature and the distortion of the magnetic field by the plasma stream. We have included the anisotropy of the thermal conductivity of a plasma in a magnetic field in an MHD approximation to investigate problems in solar physics in [9]. Taking into account this anisotropy showed that the prolonged existence of X-ray bright points in the coronal magnetic field is possible.

Here, we continue our previous study [6] of postflare loop formation. To simulate the filling of a loop with plasma and heating of this loop, we used an improved version of the PERESVET code that can take into account anisotropy of the heat conduction. This code retains all important features of the older version, viz., allowances for compressibility, the use of dissipation terms, the use of the method of characteristics to impose boundary conditions, automatic multi-level division of the time step in places with large gradients, use of a fully implicit scheme with an iterative solution technique, etc.

2. THE NUMERICAL SIMULATIONS

In our numerical simulations, we solved the following non-dimensional set of MHD equations in the XY plane:

$$\frac{\partial \mathbf{B}}{\partial t} = \mathrm{curl}\,(\mathbf{V} \times \mathbf{B}) - \frac{1}{\mathrm{Re}_m}\mathrm{curl}\left(\frac{\sigma_0}{\sigma}\mathrm{curl}\mathbf{B}\right), \quad (1)$$

$$\frac{\partial \rho}{\partial t} = -\mathrm{div}(\mathbf{V}\rho), \quad (2)$$

$$\frac{\partial \mathbf{V}}{\partial t} = -(\mathbf{V}, \nabla)\mathbf{V} - \frac{\beta_0}{2\rho}\nabla(\rho T) \quad (3)$$

$$- \frac{1}{\rho}(\mathbf{B} \times \mathrm{curl}\mathbf{B}) + \frac{1}{\mathrm{Re}\rho}\Delta\mathbf{V} + G_g\mathbf{G},$$

$$\frac{\partial T}{\partial t} = -(\mathbf{V}, \nabla)T - (\gamma - 1)T\mathrm{div}\mathbf{V} \quad (4)$$

$$+ (\gamma - 1)\frac{2\sigma_0}{\mathrm{Re}_m \sigma \beta_0 \rho}(\mathrm{curl}\mathbf{B})^2 - (\gamma - 1)G_q \rho L'(T)$$

$$+ \frac{\gamma - 1}{\rho}\mathrm{div}(\mathbf{e}_{||}\kappa_{dl}(\mathbf{e}_{||}, \nabla T)$$

$$+ \mathbf{e}_{\perp 1}\kappa_{\perp dl}(\mathbf{e}_{\perp 1}, \nabla T) + \mathbf{e}_{\perp 2}\kappa_{\perp dl}(\mathbf{e}_{\perp 2}, \nabla T)).$$

The size of the computational domain L_0 was chosen as the unit length. This corresponds to the typical size of a solar active region in which a flare has occurred, $L_0 \sim 10^{10}$ cm. Thus, we specify our computational domain to be $0 \leq X \leq 1, 0 \leq Y \leq 1$ in dimensionless units. The mean photospheric magnetic field in the active region B_0 is used as the unit of the magnetic field. As units of the plasma density and temperature, we choose their coronal values ρ_0 and T_0, assuming they are spatially uniform at the initial time. As units of velocity, time, current density, and dipole magnetic moment, we use the Alfvén speed $V_0 = V_A = B_0/\sqrt{4\pi\rho_0}$ and the quantities $t_0 = L_0/V_0$, $j_0 = cB_0/(4\pi L_0)$, and $M_0 = B_0 L_0^3$, respectively.

If the mean magnetic field in the active region is $B_0 = 300$ G, and the density of the hydrogen plasma over the active region is $\sim 10^8$ cm^{-3}, the Alfvén speed is $V_A = 6 \times 10^9$ cm/s. We assume the size of the active region to be $L = 10^{10}$ cm and obtain the value ~ 2 s for the unit of time.

In Eqs. (1)–(4), $\gamma = 5/3$ is the specific-heat ratio; $\mathrm{Re}_m = V_0 L_0/\nu_{m0}$ is the magnetic Reynolds number; $\nu_{m0} = \mathrm{C}^2/4\pi\sigma_0$ is the magnetic viscosity, based on the coronal-plasma conductivity σ_0 at temperature T_0; $\sigma/\sigma_0 = T^{3/2}$; $\beta = 8\pi n_0 k T_0/B_0^2$ ($n_0 = \rho_0/m_i$, where m_i is the ion mass); $\mathrm{Re} = L_0 V_0/\eta$ is the Reynolds number; η is the viscosity; $G_q = L(T_0 T)\rho_0 t_0/T_0$; $L(T_0 T)$ is the emission function for ionization equilibrium of the solar corona [10]; $L'(T) = L(T_0 T)/L(T_0)$ is the dimensionless emission function; $\mathbf{e}_{||}$, $\mathbf{e}_{\perp_1}$, $\mathbf{e}_{\perp 2}$ are orthogonal unit vectors directed along and perpendicular to the magnetic field; $\kappa_{dl} = \kappa/(\Pi\kappa_0)$ is the dimensionless thermal conductivity in the direction of the magnetic field; $\Pi = \rho_0 L_0 V_0/\kappa_0$ is the Péclet number; κ_0 is the thermal conductivity for a coronal temperature of T_0; κ is the thermal conductivity; $\kappa/\kappa_0 = T^{5/2}$; $\kappa_{\perp dl} = [(\kappa\kappa_0^{-1}\Pi^{-1})(\kappa_B\kappa_{0B}^{-1}\Pi_B^{-1})]/[(\kappa\kappa_0^{-1}\Pi^{-1}) + (\kappa_B\kappa_{0B}^{-1}\Pi_B^{-1})]$ is the dimensionless thermal conductivity perpendiular to the magnetic field; $\Pi_B = \rho_0 L_0 V_0/\kappa_{0B}$ is the Péclet number, based on the thermal conductivity across a strong magnetic field (with the cyclotron radius much smaller than the mean free path); the thermal conductivity denoted κ_{0B} corresponds to the temperature T_0, plasma density ρ_0, and magnetic field B_0; $\kappa_B/\kappa_{0B} = \rho^2 B^{-2} T^{-1/2}$; and $G_g\mathbf{G}$ is the dimensionless gravitational acceleration.

The computations were mainly conducted for the following set of dimensionless parameters: $\mathrm{Re}_m = 10^5$, $\mathrm{Re} = 10^4$, $\beta = 10^{-6}$, $\Pi = 100$, and $\Pi_B = 10^8$. These correspond to the principle of limited simulations [11]; i.e., parameters that are much larger than unity in the corona (such as Re_m) are assumed to be much larger than unity in the numerical simulations, although they can differ in order of magnitude. Test computations were carried out for $\mathrm{Re} = 10$. The correspondence of the chosen Reynolds numbers to the coronal conditions will be discussed below.

The magnetic field was specified by four subphotospheric vertical dipoles with alternating polarities. Their moments and coordinates are $\mu_1 = 26.25$, $X = 0$, $Y = -1.5$; $\mu_2 = -48.125$, $X = 0.25$, $Y = -1.5$; $\mu_3 = 48.125$, $X = 0.75$, $Y = -1.5$; $\mu_4 =$

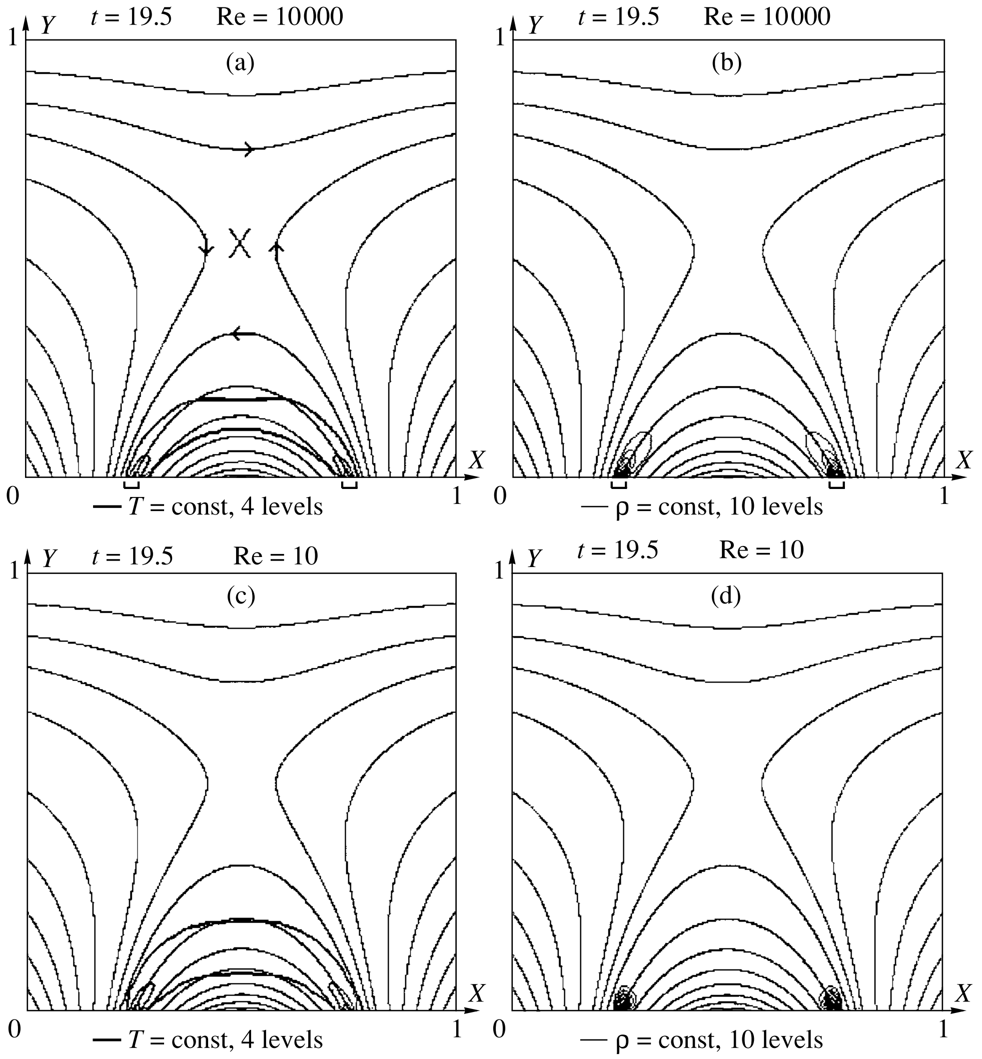

Fig. 1. Magnetic field lines with contours of (a, c) temperature and (b, d) density superposed, for Reynolds numbers of (a, b) 1000 and (c, d) 10. For simplicity, the magnetic-field direction and zero point for these and the following diagrams are shown only in the first diagram.

-26.25, $X = 1$, and $Y = -1.5$. If this arrangement is chosen for the dipoles, the magnetic field has a neutral point in the computational plane. It was shown earlier in [12] that a vertical current sheet can form near this neutral point as photospheric disturbances arise, with the flare energy accumulated in the magnetic field of this sheet. The decay of the vertical current sheet restores a potential magnetic-field configuration with a neutral point. Emerging chromospheric sources should be situated at the feet of the arched magnetic field lines.

The chromospheric plasma sources were centered at the points $X = 0.25$ and $X = 0.75$ in the photosphere. These points were on the same magnetic-field line of the unperturbed field and were adjacent to the separatrices during the formation of the current sheet. According to the electrodynamic flare model, chromospheric plasma near these points should be heated during the flare by a stream of electrons accelerated by the field-aligned currents. The parameters of the chromospheric source plasma were specified to be $T = 10$ and $\rho = 2000$.

3. SIMULATION RESULTS

Figure 1a shows magnetic-field lines and their directions. The magnetic field contains a neutral point, marked with a cross. The same figure depicts isotherms for time 19.5. The level increment is equal to the source temperature divided by the number of levels. The calculations were done for a Reynolds number of 10 000. Contours of the plasma density for the same time are given in Fig. 1b. These data show that, 40 s after the onset of chromospheric evaporation, the filling of the tube with plasma is only at an

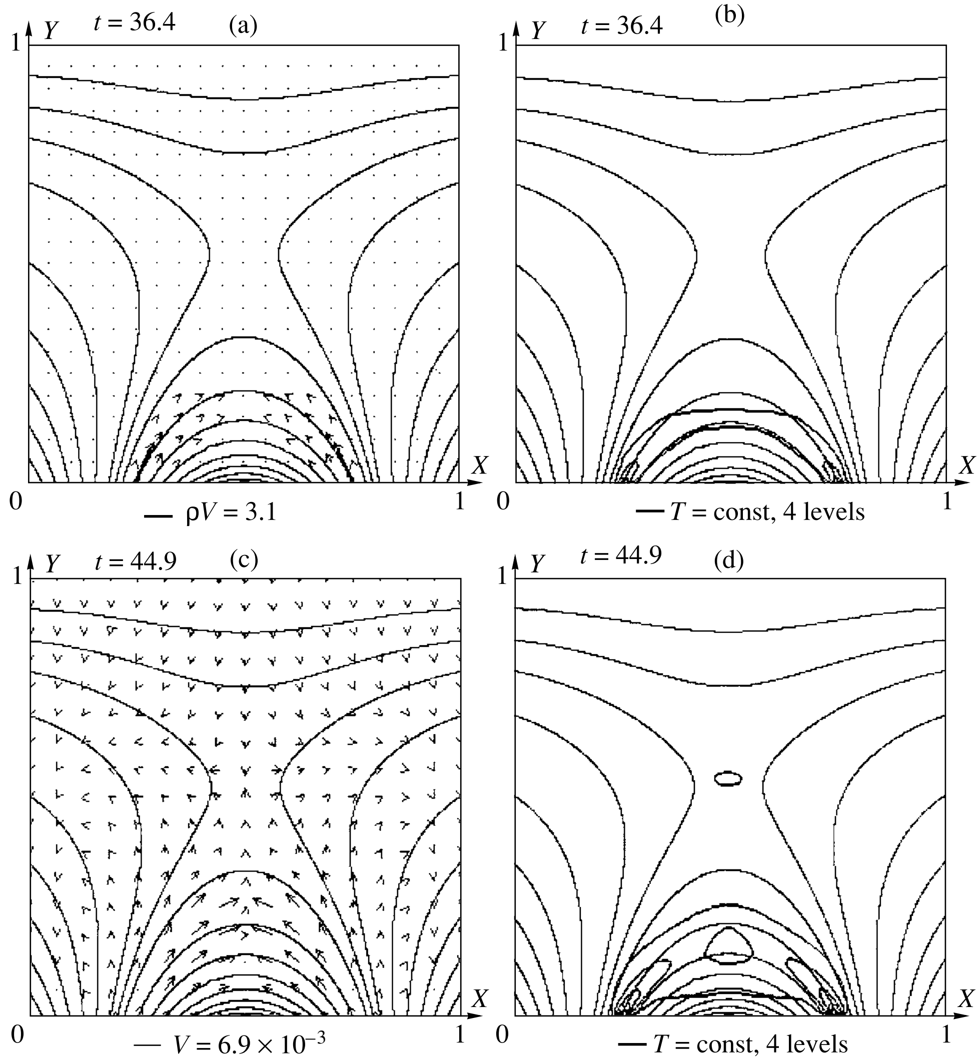

Fig. 2. Magnetic field lines with (a, c) plasma-stream vectors and (b, d) isotherms superposed, at times (a, b) $t = 36.4$ and (c, d) $t = 44.9$. Plasma heating near the neutral point and in the loop at its apex is visible.

early stage, while the plasma in the tube is completely heated, due to its high thermal conductivity. The temperature at the loop apex has more than doubled. Due to the low thermal conductivity of the plasma transverse to the field, the heat propagates almost exactly along the magnetic-field lines.

Figures 1c and 1d present computational results for the same time but a Reynolds number of 10, i.e., for a higher viscosity. The temperature behaves similarly, but the plasma stream propagates much more slowly, because viscosity leads to a strong interaction with the surrounding coronal plasma.

Figure 2 shows the computational results for a later time and Re = 10 000. We can see in Fig. 2a that, by $t = 36$, the flux of plasma ρV has approached the top of the loop. At this time, the heating of the loop continues (Fig. 2b). In Fig. 2c, the pattern of the magnetic-field lines is superposed with plasma-velocity vectors. At the loop apex, the plasma streams ascending from the two feet have started interacting. The oppositely directed, interacting plasma streams are accelerated by the pressure gradient during their motion along the field lines. Their speeds reach dimensionless values of 6×10^{-3} near the apex, before the onset of their interaction. For a mean active-region magnetic field of $B_0 = 300$ G, the corresponding dimensional speed is about 3×10^7 cm/s. This is the maximum speed that can be reached by a small fraction of the stream at its gently sloping front. The bulk speed of the stream is a factor of a few smaller. The quantity $V = 3 \times 10^7$ sm/s is almost an order of magnitude larger than the mean velocities observed in loops, which correspond to the densest portion of the stream.

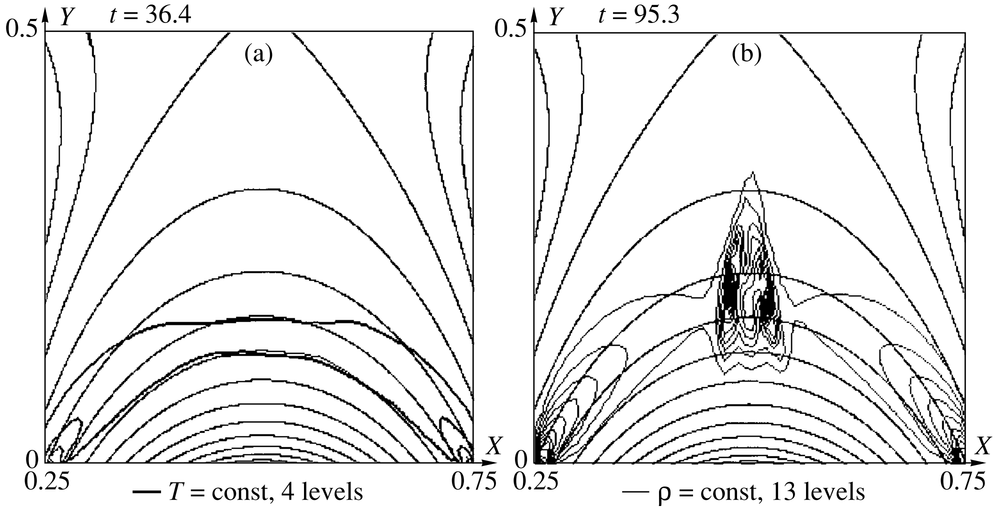

Fig. 3. Magnetic field lines on an enlarged scale, with (a) isotherms and (b) density contours superposed.

Another characteristic feature of the loop-plasma behavior is the appearance of small speeds over the entire computational domain (Fig. 2b). In the region adjoining the photosphere, the velocities are mainly aligned with the field, due to the weak viscous entrainment of the coronal plasma, which is manifest even at a Reynolds number of 10 000. In the upper portion of the domain, the velocity vectors are perpendicular to the magnetic field, which corresponds to the propagation of MHD disturbances. The vectors are directed toward the neutral point above and below this point, and away from it on the right and left. This type of flow pattern is typical of field-line reconnection: the plasma flows into the region of the neutral point from two opposite sides and flows out in the two other directions after reconnection of oppositely directed magnetic-field lines [13]. This pattern confirms earlier conclusions that the magnetic fields can focus disturbances near neutral points and that MHD disturbances can give rise to reconnection. This reconnection heats the plasma near the neutral point, as can be seen in Fig. 2d. Another center of plasma heating that can be distinguished in this figure is in the loop, at its apex, where the energy of the two colliding plasma streams is dissipated. To our knowledge, this additional mechanism for plasma heating at the loop apex has not been considered previously.

Figures 3a and 3b show isotherms and density contours on an enlarged scale. The situation reflected by Fig. 3b corresponds to a fairly long injection, with the streams efficiently interacting at the loop apex. The plasma density there has increased by more than an order of magnitude; however, magnetic pressure still confines the plasma, preventing it from substantially distorting the shape of the field lines.

4. DISCUSSION

Using an MHD approximation to investigate processes occuring in cosmic plasmas has proved to be very fruitful. Early studies applying MHD techniques to the solar corona have been reviewed in a number of monographs, in particular, that of Priest [14]. A more detailed investigation of several coronal processes is impossible if the MHD equations are used in their simplest form; i.e., without taking into account the anisotropy of the thermal conductivity and the viscosity of the plasma in the magnetic field, the Hall conductivity, etc.. The introduction of thermal-conductivity anisotropy into the PERESVET code enabled us to simulate X-ray coronal bright points, and show that they can persist over long times near magnetic-field neutral lines [9].

Previously published MHD calculations of post-flare loop dynamics did not consider anisotropy of the thermal conductivity and assumed a $k \sim T^{5/2}$ law. These calculations showed filling of a loop with plasma from flare-induced chromospheric sources. However, these simulations suggested that the plasma that flows from the sources should lose heat to ambient space and rapidly cool. As a result, the temperature of the loop should differ only slightly from the coronal temperature. Taking into account the thermal-conductivity, anisotropy demonstrated heating of a post-flare loop by hot photospheric sources. The propagation of the heat energy substantially outstrips the propagation of material from sources of chromospheric evaporation. The rapid heat propagation along the loop reduces the thermal gradient in this direction so that the pressure gradient is also reduced. For this reason, when the thermal-conductivity anisotropy is taken

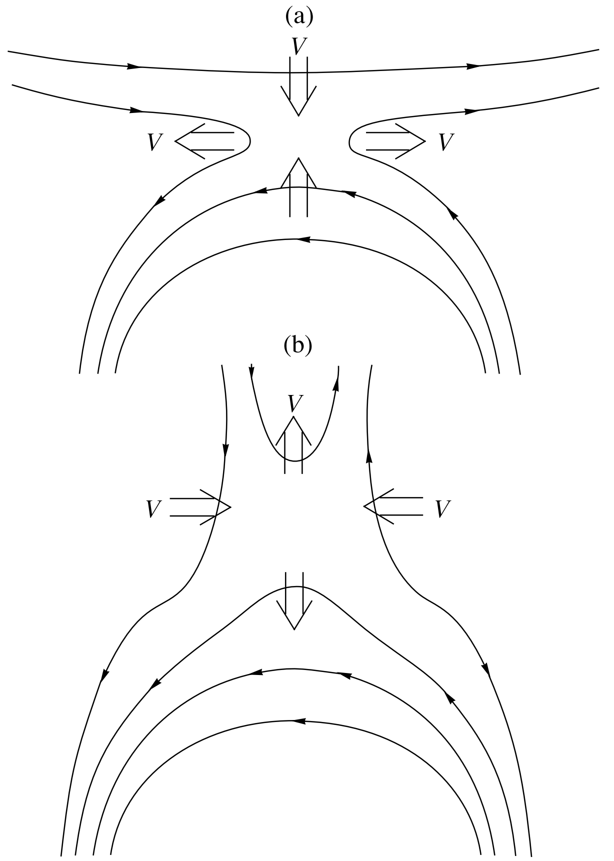

Fig. 4. Patterns of reconnection over the loop apex for a (a) horizontal and (b) vertical current sheet. The loop can be heated via reconnection only in the latter case.

into account, the acceleration of the plasma flowing from the source proves to be much smaller than in the case of isotropic heat conduction [6], and, for the same field parameters and source intensity, the filling of the flux tube is about a factor of two slower.

The Hall field $\mathbf{j} \times \mathbf{B}/nec$ in the Ohm law is zero everywhere in the corona except in current sheets, where it nevertheless remains small compared to the Lorentz field $\mathbf{V} \times \mathbf{B}/c$ and does not substantially affect the plasma dynamics. However, including the effect of the Hall field in coronal current sheets and the geomagnetic tail makes it possible to explain the main current systems arising in the corona during flares [5] and in the magnetosphere during substorms [15]. In our calculations, $\mathbf{j} \times \mathbf{B}$ is small, and the Hall term can be neglected.

As has been done in all previously published studies, we have neglected the anisotropy of the viscosity due to the magnetic field. This distorts our results to some extent. The ratio of the effects of convection and viscosity is determined by the Reynolds number $\mathrm{Re} = LV/V_{Ti}\lambda$. In the solar corona, the mean free path is comparable to the size of a post-flare loop; therefore, $\mathrm{Re} \sim 1$. In an MHD calculation of a field-aligned plasma stream, a Reynolds number $\mathrm{Re} \sim 1$ adequately characterizes the interaction between the stream and the coronal plasma in the flux tube channeling the stream. However, there is a strong viscous entrainment of the ambient plasma in this case, as if there were no magnetic field. At the same time, in the presence of a magnetic field, the characteristic transport length is the ion Larmor radius ρ_i rather than the ion mean free path. This means that the cross-field viscosity of the magnetized plasma is reduced by the factor $(\lambda/\rho_i)^2$. Since $\rho_i \ll \lambda$ in an active region, the entrainment of the plasma ambient to the loop is not significant there.

Indeed, test computations showed that, at small Reynolds numbers ($\mathrm{Re} \sim 10$), the ambient plasma is entrained during the motion, and the stream is strongly decelerated. This does not correspond to the conditions for filling a post-flare loop. To weaken this entrainment of the ambient plasma, we assumed $\mathrm{Re} = 10\,000$ in our computations. Under these conditions, the entrainment of the coronal plasma surrounding the flux tube becomes insignificant, as should be the case in the corona. In the interaction of the plasma streaming from the chromospheric source with plasma already occupying the flux tube, viscosity affects only the width of the diffusive boundary layer. This interaction in the loop does not play an important role in the stream dynamics, since the plasma density assumed for the chromospheric source exceeds the coronal density by a factor of 2000. The effect of the Reynolds number on the dynamics of the flow in the loop is illustrated in Fig. 1, which shows the initial stage ($t = 19.5$) of plasma injection into the loop, when the plasma streams from the sources have not yet started to interact at the loop apex, for $\mathrm{Re} = 10\,000$ and $\mathrm{Re} = 10$. The heating of the loop is essentially the same in the two cases, but the viscous interaction of the streams with the ambient coronal plasma results in a distinct deceleration of the stream at the lower Reynolds number.

When $\mathrm{Re} \gg 1$, the entrainment of the coronal plasma in motion outside the magnetic tube is weak, but the viscous interaction of the two oppositely directed streams at the loop apex is strong, and the kinetic energy of the streams is dissipated within a very narrow region. In this region, there is a sharp, shock-like increase in the plasma density and temperature. Under natural conditions, the ion mean free path is long, the streams should not interact so strongly, and the width of the interaction region should be of the order of the mean free path.

This model for a post-flare loop with hot chromospheric sources (kernels) at both feet, which originate after magnetic reconnection in the corona, is in agreement with observations in OV and FeXXI lines

[16]. Qiu *et al.* [17] reported oppositely directed plasma streams with speeds in excess of 30 km/s.

It has been noted in a number of studies that many loops have non-uniform temperature distributions. Hotter plasma is located in the higher part of the loop, while its interior is much cooler [1, 18]. The extra heating of the outer part of the loop can naturally be explained by the release of heat due to magnetic reconnection. If the magnetic field at the top of the loop is directed opposite to the field over the loop, a neutral line should exist at the boundary, making reconnection possible. A schematic of the reconnection and plasma streams in this situation is given in Fig. 4a. The plasma is heated in the process of reconnection, and magnetic-field lines disappear in the outer portion of the loop [5]. The closed field lines of the loop become open, the low thermal conductivity of the plasma prevents the diffusion of heat into the loop, and heat propagates along field lines that no longer belong to the arch.

A different situation arises if a vertical current sheet was present above the arch (Fig. 4b) and its breakdown resulted in a flare [12]. Reconnection in the vertical current sheet accelerates the plasma both upward and downward along the sheet via the magnetic-tension force. The upward accelerated plasma is ejected from the corona into interplanetary space, while the downward directed stream of hot plasma carries fresh magnetic-field lines that have reconnected in the current sheet into the arch. In this case, the plasma heated most strongly during the reconnection should be in the outer part of the loop, with cooler plasma inside the loop.

5. CONCLUSION

(1) In our numerical MHD simulations of the formation of a post-flare loop, we have taken into account for the first time anisotropy of the thermal conductivity of plasmas in magnetic fields.

(2) The heating of the loop is substantially more rapid than its filling with chromospheric plasma from localized hot sources.

(3) The plasma streams from two chromospheric sources interacting at the loop apex produce additional heating there.

(4) The frequently observed electron-temperature maximum in the outer part of loops seems to be related to reconnection process in vertical current sheets over loops.

ACKNOWLEDGMENTS

We are grateful to L.V. Surkova for important technical assistance. This work was supported by the Russian Foundation for Basic Research (project nos. 01-02-16186 and 00-01-00091).

REFERENCES

1. B. Schmieder, P. Heinzel, L. van Dril-Gesztelyi, *et al.*, Sol. Phys. **165**, 303 (1996).
2. Z. Svestka, *Solar Flare* (D. Reidel, Dordrecht, 1976).
3. G. H. Fisher, R. C. Canfield, and A. N. McClimon, Astrophys. J. **289**, 414 (1985).
4. J. I. Culhane, A. T. Phillips, M. Inda-Koide, *et al.*, Sol. Phys. **153**, 307 (1994).
5. A. I. Podgorny and I. M. Podgorny, Sol. Phys. **139**, 125 (1992).
6. A. I. Podgorny and I. M. Podgorny, Astron. Zh. **75**, 132 (1998) [Astron. Rep. **42**, 116 (1998)].
7. S. Tsuneta, Astron. Soc. Pac. Conf. Ser. **46**, 239 (1993).
8. A. I. Podgorny and I. M. Podgorny, Astron. Zh. **76**, 696 (1999) [Astron. Rep. **43**, 608 (1999)].
9. A. I. Podgorny and I. M. Podgorny, Astron. Zh. **77**, 467 (2000) [Astron. Rep. **44**, 407 (2000)].
10. D. P. Cox and W. H. Tucker, Astrophys. J. **157**, 1157 (1969).
11. I. M. Podgorny, Fundam. Cosm. Phys. **4**, 1 (1978).
12. A. I. Podgorny and I. M. Podgorny, Astron. Zh. **78**, 71 (2001) [Astron. Rep. **45**, 60 (2001)].
13. A. I. Podgorny, Sol. Phys. **156**, 41 (1995).
14. E. Prist, *Solar Magnetohydrodynamics* (D. Reidel, Dordrecht, 1985).
15. A. I. Podgorny, I. M. Podgorny, and S. Minami, J. Geomagn. Geoelectr. **49**, 1099 (1997).
16. L. C. Bagala, C. H. Mandrini, and M. G. Rovira, Sol. Phys. **161**, 103 (1995).
17. J. Qiu, H. Wang, J. Chae, and P. R. Goode, Sol. Phys. **190**, 153 (1999).
18. T. G. Forbos, Philos. Trans. R. Soc. London, Ser. A **358**, 711 (2000).

Translated by A. Getling

Astronomy Reports, Vol. 46, No. 1, 2002, pp. 74–82. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 1, 2002, pp. 81–89.
Original Russian Text Copyright © 2002 by Akulenko, Kumakshev, Markov, Rykhlova.

A Model for the Polar Motion of the Deformable Earth Adequate for Astrometric Data

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

[1]*Institute for Problems in Mechanics,*
Russian Academy of Sciences,
pr. Vernadskogo 101, Moscow, 117526 Russia
[2]*Moscow Institute of Aviation,*
Volokolamskoe sh. 4, Moscow, 125080 Russia
[3]*Institute of Astronomy,*
Russian Academy of Sciences,
Pyatnitskaya ul. 48, Moscow, 109017 Russia

Received December 16, 2000

Abstract—Refined analytical expressions for the frequencies corresponding to the Chandler motion of the pole and the diurnal rotation of the deformable Earth are derived. Numerical estimates of the period and amplitude of the polar oscillations are presented. The trajectory of the Chandler polar motion derived via numerical modeling is in qualitative and quantitative agreement with experimental data from the International Earth Rotation Service (IERS). An evolutionary model describing slow variations in the Earth's rotation parameters under the action of the dissipative moments of the tidal gravitational forces on time scales considerably longer than the precession period of the Earth's axis is constructed. The axis of the Earth's figure tends to approach the angular momentum vector of the proper rotation.

1. INTRODUCTION

Dynamical theories of the Earth's rotation with respect to its center of mass have been the basis of many astrometric investigations [1–7]. The modern level of accuracy in the Earth rotation parameters [4, 7] decisively demands the exclusion of a number of simplifications from the commonly accepted theory. High-accuracy observations, measurements, and data on the trajectory of the polar motion reveal quite complex dynamical processes occuring in the Earth–Moon–Sun system, which affect the Earth's rotation [1–7]. Various types of physical perturbations—both gravitational and geophysical—can contribute to this phenomenon.

The development of an adequate theoretical model for the rotational–translational motion of a planet is one of the basic problems of astrometry. Such a mathematical model must satisfactorily describe the real trajectory of the pole (position of the angular-velocity vector) in some convenient coordinate system fixed to the Earth over the considerable observable time intervals of the order of 10–100 yr. At the preliminary stages of constructing such a model, it seems reasonable to identify and take into account only basic factors associated with celestial mechanics (gravitational perturbations) and geophysics (the deformable character of the mantle). In subsequent stages, this model can be made more complex and refined in order to explain small-scale processes, which usually display irregular behavior, to obtain better agreement with the observational data.

Along with analysis of rapid processes, it is very important and fruitful from both the astrometric and geophysical points of view to study the evolution of the Earth's rotation and the translational–rotational motion of the Earth–Moon system in the solar gravitational field over long time intervals of the order of 10^5–10^7 yr. Mathematical modeling and an asymptotic qualitative analysis of these slow evolutionary processes can be carried out using a simplified model in which the Earth has a viscoelastic mantle with a solid core at its center.

We were the first to obtain well founded analytical expressions for the coordinates of the Earth's pole, which represent a contribution to the development of a high-accuracy theory for the Chandler polar motion. These expressions yield good agreement between the results of mathematical modeling and astrometric observations obtained over many years. We have derived and analyzed an evolutionary model for the rotational motions of the deformable Earth, in particular, the very slow evolution of the angle between the figure axis and the angular momemtum.

2. INFLUENCE OF ELASTIC DEFORMATIONS ON THE EARTH'S INERTIA TENSOR

The influence of the mantle's elasticity on the Earth's rotation about its center of mass has been studied in [1–6] and other works. Because of the importance of this problem for the development of a mathematical model describing the motion of the Earth–Moon system, we present the basic assumptions connected with refining the inertia tensor for a rotating deformable Earth and calculating the angular momemtum vector and its time derivative.

The Cartesian coordinate system C_2x_i ($i = 1, 2, 3$) rigidly fixed to the solid core of the undeformed planet is introduced in the standard way (Fig. 1). The principal axes of the inertia tensor can be conveniently taken to be the corresponding coordinate axes, and the center of mass, to be the point C_2. Next, the coordinate system $C_2'x_i'$ ($i = 1, 2, 3$) for the deformed state is introduced by translating the coordinate origin from C_2 to C_2' (the center of mass taking into account small deformations), with a corresponding parallel translation of the axes. The elastic deformations $\mathbf{u}$ and $\mathbf{u}_*$ in the systems C_2x_i and $C_2'x_i'$ are related via the expression $\mathbf{u}_* = \mathbf{u} - \mathbf{u}_C$, where $\mathbf{u}_C$ is the displacement of the center of mass with respect to the core. To simplify the calculations, the mantle is assumed to be uniform and isotropic. The desired vector $\mathbf{u}$ obeys the Euler–Cauchy equation of state [8] and the boundary conditions at the Earth's surface P and core surface P_0:

$$\Delta\mathbf{u} + \frac{1}{1-2\nu}\nabla(\nabla, \mathbf{u}) \tag{1}$$

$$+\frac{\rho}{\mu}\mathbf{\Phi} = 0, \ \mathbf{n}\cdot\sigma_n|_P = 0,$$

$$\mathbf{u}|_{P_0} = 0.$$

Here, Δ is the Laplace operator, ∇ the Hamilton operator, ρ the density, ν the Poisson ratio, μ the rigidity modulus, $\mathbf{\Phi}$ the inertia force per unit mass, $\mathbf{n}$ a unit vector perpendicular to P, and σ_n the stress tensor. Further, we will study the problem (1) in a quasi-static approximation ($\ddot{\mathbf{u}} = \dot{\mathbf{u}} \equiv 0$) [5, 8, 9].

It is important to note that the term in the function $\mathbf{\Phi}$ containing $\dot{\boldsymbol{\omega}}$ (where $\boldsymbol{\omega}$ is the angular-velocity vector of the Earth's rotation) can also be neglected. This is justified by the closeness of the vector $\boldsymbol{\omega}$ to the principal axis of inertia—the axis of the Earth's figure. Let us present the required function in the form $\mathbf{u} = \mathbf{u}_0(\mathbf{r}) + \mathbf{u}_*(\mathbf{r}, t)$, where $\mathbf{u}_0$ is a quasi-static displacement (a static equatorial bulge) and $\mathbf{u}_*$ are the deformations caused by the tidal gravitational forces of the Moon and Sun. The function $\mathbf{u}_0$ is determined at the first stage of refining the inertia tensor of the deformable Earth when constructing a model for the polar motion on comparatively short time intervals. This function is the solution of the boundary problem (1) when $\mathbf{\Phi} = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ and can be represented as an expansion in the numerical parameter $\rho\omega^2R^2/\mu$, where R is a characteristic linear size (the Earth's radius). The additional term $\mathbf{u}_*$, which characterizes dissipative tidal moments, is estimated in a similar way taking into account astrometric data.

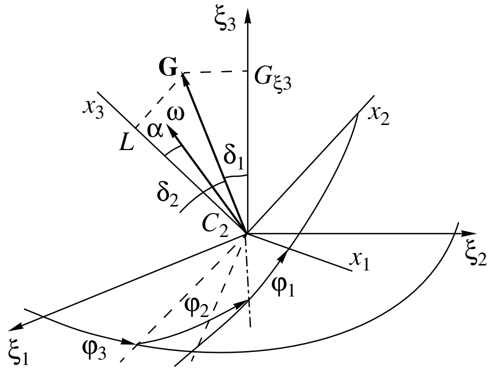

Fig. 1. Mutual orientation of fixed and inertial coordinate systems; definition of the basic characteristics of the Earth's rotation.

Let us calculate the angular momemtum vector of the Earth in the deformed state $\mathbf{G}$ and its derivative [8]:

$$\mathbf{G} = \int_\Omega (\mathbf{r} + \mathbf{u}) \times \mathbf{v}\rho dV, \tag{2}$$

$$\dot{\mathbf{G}} = \int_\Omega (\mathbf{r} + \mathbf{u}) \times (\mathbf{w} + \boldsymbol{\omega} \times \mathbf{v})\rho dV.$$

Here, dV is a volume element and $\mathbf{v}$ and $\mathbf{w}$ are the velocity and acceleration, calculated using the kinematics laws for the rotating coordinate system. The region Ω contains a rigid core, where $\mathbf{u} \equiv 0$, and a deformable mantle, where the vector $\mathbf{u}$ is determined by (1). The expressions for $\mathbf{G}$ and $\dot{\mathbf{G}}$ (2) can be simplified by neglecting terms quadratic in $\mathbf{u}$. As a result, we can obtain an expression including the principal part (the undeformed planet) and small corrections due to the displacements $\mathbf{u}$ and their derivatives with respect to t.

3. CONSTRUCTION OF THE INTERMEDIATE TRAJECTORY IN THE THEORY OF ROTATION OF THE DEFORMABLE EARTH

In our investigation of the Chandler motion of the poles and their evolution under the action of perturbing moments, we shall first consider the intermediate

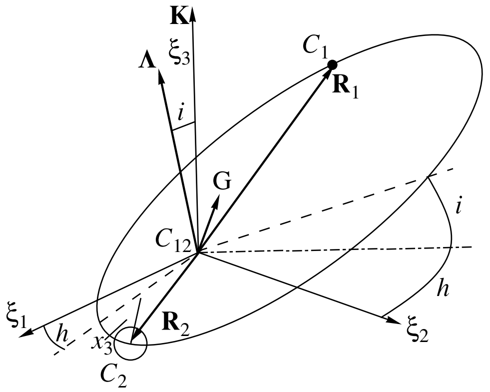

Fig. 2. Mutual motion of the Earth and Moon; definition of the inertial coordinate system and the parameters of the orbital motion.

model [5, 6]. We will assume that a deformable planet (the Earth) and a point satellite (the Moon) possess a mutual translational–rotational motion about their common center of mass (barycenter). We introduce the inertial coordinate system $C_{12}\xi_1\xi_2\xi_3$, whose origin C_{12} is at the barycenter. Let $\mathbf{G}$ be the proper angular momemtum of the Earth and $\mathbf{\Lambda}$ the total orbital angular momemtum of the centers of mass of the Moon C_1 and Earth C_2. Then, the angular momentum of the entire closed system $\boldsymbol{K} = G + \Lambda$ without taking into account external perturbations is fixed in the inertial reference frame and coincident with the $C_{12}\xi_3$ axis (Fig. 2). The radius vectors $\mathbf{R}_1$ and $\mathbf{R}_2$ of the points C_1 and C_2 in the coordinate system $C_{12}\xi_1\xi_2\xi_3$ are given in the form

$$\mathbf{R}_j = R_j\mathbf{R}_j^0, \quad R_j = c_j^* R_{21}, \tag{3}$$
$$\mathbf{R}_2^0 = -\mathbf{R}_1^0, \quad j = 1, 2,$$
$$\mathbf{R}_j^0 = \pm\Gamma_3(h)\Gamma_1(i)(\cos\vartheta, \sin\vartheta, 0)^T$$
$$c_1^* = \frac{m_2}{m}, \quad c_2^* = \frac{m_1}{m}, \quad m = m_1 + m_2$$
$$\Gamma_1(i) = \mathrm{diag}(1, \Pi_2(i)),$$
$$\Gamma_3(h) = \mathrm{diag}(\Pi_2(h), 1).$$

The matrices $\Gamma_{1,3}$ in (3) have a block-diagonal form, and Π_2 is the matrix of plane-rotation. The angular variables h, i, and ϑ are the longitude of the ascending node, inclination, and true anomaly of the orbit, respectively. The radius vector $\mathbf{R}_{21}$ between the centers of mass of the Earth and Moon (i.e., between the points C_2 and C_1) has the form $\mathbf{R}_{21} = R_{21}\mathbf{R}_{21}^0$, so that $\mathbf{R}_{21}^0 = \mathbf{R}_1^0$. Below, we shall use the numerical values $R_{21} = 384.4 \times 10^6$ m, the mass of the Moon $m_1 = 7.36 \times 10^{22}$ kg, and the mass of the Earth $m_2 = 5.98 \times 10^{24}$ kg.

The Cartesian coordinate system $C_2x_1x_2x_3$ with its axes directed along the principal axes of inertia at the center of gravity A, B, and C is rigidly fixed to the solid core of the planet. The basis vector $\mathbf{R}_{21}^0$ in this coordinate system is determined by the relation

$$S^{-1}\mathbf{R}_{21}^0 = (\gamma_1, \gamma_2, \gamma_3)^T, \tag{4}$$
$$S^{-1} = \Gamma_3^{-1}(\varphi_1)\Gamma_1^{-1}(\delta_2)\Gamma_3^{-1}(\varphi_2)\Gamma_1^{-1}(\delta_1)\Gamma_3^{-1}(\varphi_3).$$

Here, the orthogonal matrix $S = S(t)$ $(S^{-1} = S^T)$ describes the transformation from the fixed to the inertial axes and is expressed in terms of the canonical variables of Andoyer [5, 6]—the angular momenta L, G, $G_{\xi 3}$ and the angles φ_1, φ_2, φ_3. The angles δ_1 and δ_2 are determined by the relations (Fig. 1)

$$\cos\delta_1 = \frac{G_{\xi 3}}{G}, \quad \cos\delta_2 = \frac{L}{G}, \tag{5}$$

where G is the absolute value of the Earth's angular momemtum and $G_{\xi 3}$ is the projection of G onto the inertial axis ξ_3.

The mutual orbital motion of the centers of mass C_1 and C_2 is described by the Delone canonical variables Λ, H, ϑ, and h, where $\Lambda = |\mathbf{\Lambda}|$, and $H = \Lambda\cos i$ is the projection of $\mathbf{\Lambda}$ onto the $C_{12}\xi_3$ axis, i.e., the axis of the total angular momemtum $\mathbf{K}$.

After a series of straightforward manipulations and averaging over the rapid variables φ_2 and ϑ, the Routh functional R^* in the intermediate problem for a deformable Earth can be reduced to the form (with accuracy to within an unimportant constant) [5, 6]

$$R^* = \frac{G^2 - L^2}{2}\left(\frac{\sin^2\varphi_1}{A^*} + \frac{\cos^2\varphi_1}{B^*}\right) \tag{6}$$
$$+\frac{L^2}{2C^*} + \mathrm{const}\,.$$

Here, A^*, B^*, and C^* are the effective principal central moments of inertia taking into account deformations of the frozen Earth due to its complex motion—its proper rotation and motion with respect to the barycenter C_{12}. We shall use the numerical values $A^* = 8.0912 \times 10^{37}$ kg m^2, $B^* = 8.0914 \times 10^{37}$ kg m^2, and $C^* = 8.1100 \times 10^{37}$ kg m^2. These quantities can be expressed

$$A^*, B^*, C^* \tag{7}$$
$$= J_{jj}[\mathbf{u}]I_2^2(I_2^2 + 3\mu_1R_{21}^{-3}dJ_{jj}^2[\mathbf{u}])^{-1}\big|_{j=1,2,3},$$
$$d = \cos^2 h^* + \cos^2 i\sin^2 h^* + \cos^2\delta_1(\sin^2 h^*$$
$$+ \cos^2 i\cos^2 h^*) + \sin^2 i\sin^2\delta_1,$$
$$h^* = \varphi_3 - h, \quad \mu_1 = fm_1, \quad I_2 = G.$$

Here, f is the gravitational constant, and $J_{jj}[\mathbf{u}]$ are functionals of the elastic displacements $\mathbf{u}$ of the particles of the viscoelastic mantle (which are, in fact, components of the principal central inertia tensor for the deformed Earth in the fixed coordinate system). When $\mathbf{u} = 0$, we obtain $J_{jj}[0] =$

$\mathrm{diag}(A, B, C)$. Therefore, in accordance with (7), we can treat the Earth–Moon system nearly as a binary planet, but the contribution from the motion about its barycenter is relatively small ($\sim 5\%$).

The structure of (6) coincides with the usual expression for the Routh function for a rigid body in Andoyer variables [5, 6]. This is the basis for introducing the action–angle variables I_j, w_j $(j = 1, 2, 3)$, which are very convenient for asymptotic methods. Note that the action–angle variables specify a Hamiltonian system of variables.

Next, we introduce the basic dynamic parameters $\varkappa$ and λ characterizing the rotational motions of the deformable Earth:

$$\varkappa^2 = \frac{C^*(A^* - B^*)}{A^*(B^* - C^*)}, \quad \lambda^2 = \varkappa^2\frac{2EC^* - I_2^2}{I_2^2 - 2EA^*}. \tag{8}$$

The constant E represents the kinetic energy integral of the intermediate motion (rotation) of the deformable Earth. We will take the angular momemtum of the Earth's rotation about its axis $I_2 = G$ to be $I_2 = 5.85 \times 10^{33}$ kg m^2 (see below).

The relation between the action–angle variables I_j, w_j $(j = 1, 2, 3)$ and the Andoyer canonical variables L, G, $G_{\xi 3}$, φ_j is not uniquely determined in terms of elliptic functions and integrals [5, 6]. In the case of the Earth's rotation, we have

$$I_1 = \frac{2I_2}{\pi\varkappa}\frac{\varkappa_*}{\chi}\left[\chi^2\Pi\left(\frac{\pi}{2}, \varkappa^2, \lambda\right) - \lambda^2\mathbf{K}(\lambda)\right], \tag{9}$$

$$\varkappa_* = \sqrt{1 + \varkappa^2}, \ \chi = \sqrt{\varkappa^2 + \lambda^2},$$

$$0 < \lambda < 1, \ I_2 = G, \ I_3 = G_{\xi 3},$$

$$w_1 = \pm\frac{\pi}{2}\frac{F(\zeta, \lambda)}{\mathbf{K}(\lambda)}, \quad \tan\zeta = -\varkappa_*^{-1}\cot\varphi_1,$$

$$w_2 = \varphi_2 \pm \frac{\varkappa_*\chi}{\varkappa}$$

$$\times\left[\Pi(\zeta, \varkappa^2, \lambda) - \Pi(\frac{\pi}{2}, \varkappa^2, \lambda)\frac{F(\zeta, \lambda)}{\mathbf{K}(\lambda)}\right],$$

$$\zeta = \pm am(\eta, \lambda), \quad \eta = \frac{2}{\pi}\mathbf{K}(\lambda)w_1,$$

$$w_3 = \varphi_3.$$

Here, $F(\zeta, \lambda)$ and $\Pi(\zeta, \varkappa^2, \lambda)$ are elliptic integrals of the first and third kinds, respectively, and $\mathbf{K}(\lambda)$ and $\Pi(\pi/2, \varkappa^2, \lambda)$ are the corresponding complete elliptic integrals.

Using the action–angle variables introduced in (9), we can write the Routh functional for the intermediate model problem and construct trajectories in the I_j, w_j phase space. This is necessary for our subsequent analysis of the perturbed motions at asymptotically large times, over which there are significant variations in the slow (osculating) variables I_j.

4. SEMIANALYTICAL TWO-FREQUENCY MODEL OF THE CHANDLER POLAR MOTION

Let us write the Routh functional R_0 averaged over φ_2 and ϑ, which corresponds to rotations of the deformed Earth without regard for the tidal moments [5, 6]:

$$R_0 = \frac{1}{2}\frac{I_2^2}{A^*}\left(1 - \frac{(C^* - A^*)}{C^*}\frac{\varkappa^2}{\chi^2}\right). \tag{10}$$

Here, $\varkappa^2$, A^*, and C^* are the functionals of $\mathbf{u}$ defined by (7) and (8), where $\mathbf{u}$ is the vector of elastic deformations describing a quasi-static equatorial bulge due to the planet's rotation. The quantity R_0 depends also on the action variables I_1 and I_2 [see (8)], but not on other phase variables. As a result, the general solution of the problem at hand takes the form

$$I_j(t) = I_j(0) = I_j^0, \ j = 1, 2, 3, \ w_3 = w_3^0, \tag{11}$$

$$w_1 = n_1 t + w_1^0, \quad w_2 = n_2 t + w_2^0,$$

$$n_1 = -\frac{\pi}{2}I_2\frac{C^* - A^*}{A^*C^*}\varkappa\varkappa_*^{-1}\chi^{-1}\mathbf{K}^{-1}(\lambda),$$

$$n_2 = \frac{I_2}{C^*}\left(1 + \frac{(C^* - A^*)}{A^*}\Pi(\frac{\pi}{2}, \varkappa^2, \lambda)\mathbf{K}^{-1}(\lambda)\right).$$

We will take the angular velocity of the Earth's rotation to be $n_2 = 7.27 \times 10^{-5}$ s^{-1}.

The quantities I_j^0 and w_j^0 in (11) are the initial values determined via the Andoyer variables at $t = 0$ by expressions (9). The phases w_1, w_2 and frequencies n_1, n_2 correspond to the Chandler polar motion and diurnal rotation of the deformable Earth, respectively. Therefore, in the intermediate motion, the deformable Earth rotates uniformly in the field of the centrifugal forces and the gravitational field of the Moon. In the approximation of a rigid planet ($\mathbf{u} \equiv 0$), we obtain regular Euler–Poinsot precession. In the case of a deformable Earth, expressions (11) also describe regular precession, but the angular velocities of the precession and proper rotation are changed by a relatively small amount due to the perturbing factors. In the case of perturbed motion taking into account the dissipative properties of the viscoelastic mantle, we obtain regular precession with parameters that vary slowly with time; i.e., there is evolution of the slow variables, which can be studied using asymptotic methods of nonlinear mechanics (see below). The resulting expressions (11) are the generating solution for the averaging method.

Estimates of λ for the Earth–Moon system show that $\lambda^2 \ll 1$ ($\lambda^2 \sim 10^{-14}$); i.e., the ratio $(I_2 - I_1)/I_2$ is small. To a first approximation, we find from the first relation of (9) that

$$\lambda^2 = \frac{2\varkappa^2(I_2 - I_1)}{I_2\varkappa_*}. \tag{12}$$

Substituting λ^2 from (12) into (11), we obtain for the frequencies n_1 and n_2 of the intermediate motion

$$\begin{aligned}\dot{w}_1 = n_1 &= -\frac{I_2}{\varkappa_*}\frac{C^*-A^*}{A^*C^*} \qquad (13)\\ &\times\left(1-\frac{2+\varkappa^2}{2\varkappa_*}\frac{I_2-I_1}{I_2}\right),\\ \dot{w}_2 = n_2 &= \frac{I_2}{C^*}\left(1+\frac{C^*-A^*}{\varkappa_* A^*}\right.\\ &\left.\times\left(1+\frac{I_2-I_1}{2\varkappa_* I_2}(2\varkappa_*-2-\varkappa^2)\right)\right).\end{aligned}$$

The approximate analytic expressions (7) for the Earth's moments of inertia taking into account its deformation and translational motion in the Earth–Moon system represent a basis for constructing a high-accuracy model of the Chandler polar motion. Using formula (13) for n_1, we obtain the angular velocity and oscillation period $T_1^* = 2\pi/n_1 \approx 430$ d; measurements of the oscillation period yield $T_1 \approx 420$–440 sidereal days [3, 4, 7].

Polar motion is defined as the angular displacement of the rotational axis in the body of the planet with respect to its figure axis C_2x_3. The components of the angular velocity vector are approximately equal to

$$\begin{aligned}\omega_x &= \lambda I_2\frac{\mathrm{cn}(\eta,\lambda)}{A^*\omega\chi}, \quad \omega_y = \lambda I_2\frac{\varkappa_*\,\mathrm{sn}(\eta,\lambda)}{B^*\omega\chi}, \qquad (14)\\ \omega_z &= \frac{I_2}{C^*}\frac{\varkappa\,\mathrm{dn}(\eta,\lambda)}{\chi}, \quad \omega = \sqrt{\omega_x^2+\omega_y^2+\omega_z^2}\\ &= \frac{I_2}{A^*\chi}\left(\lambda^2\,\mathrm{cn}(\eta,\lambda)+\left(\frac{A^*}{C^*}\right)^2\right.\\ &\left.\times\,\varkappa^2\left(1-\lambda^2\sin^2(\eta,\lambda)\right)\right)^{1/2}.\end{aligned}$$

The angular coordinates x_p and y_p can be calculated as follows:

$$\begin{aligned}x_p &= \frac{\omega_x}{\omega} \approx \frac{C^*}{A^*}\frac{\lambda}{\varkappa}\cos w_1, \qquad (15)\\ y_p &= -\frac{\omega_y}{\omega} \approx \frac{C^*}{B^*}\frac{\lambda}{\varkappa}\varkappa_*\sin w_1,\\ \cos\alpha &\approx 1-\frac{1}{2}\left(\frac{\lambda}{\varkappa}\right)^2\\ &\times C^{*2}(A^{*-2}+\varkappa_* B^{*-2})\sin^2 w_1.\end{aligned}$$

The minus sign in the expression for y_p (15) was taken in accordance with international agreement; i.e., the angular coordinate y_p is directed along the meridian 90° W [3, 4, 7]. Here, α is the angle between the angular-velocity vector and the Earth's figure axis (Fig. 1). The approximate expressions (15) have very high accuracy, since $\lambda^2 \sim 10^{-14}$ and $\lambda^2/\varkappa^2 \sim 10^{-12}$.

The curve drawn by the angular-velocity vector in the Earth's body in the tangent plane after one complete revolution is nearly closed. It is approximately an ellipse with semiaxes D_x and D_y, which are nearly equal:

$$\begin{aligned}&X_p^2D_x^{-2}+Y_p^2D_y^{-2}=1, \quad X_p = D_x\cos w_1, \qquad (16)\\ &Y_p = -D_y\sin w_1, \quad D_x = \rho\frac{C^*}{A^*}\frac{\lambda}{\varkappa},\\ &D_y = \rho\frac{C^*}{B^*}\frac{\lambda}{\varkappa}\varkappa_*, \quad \frac{D_y}{D_x} = \frac{A^*}{B^*}\varkappa_* > 1.\end{aligned}$$

Here, $\rho = 6380$ km is the Earth's radius. It follows from (15) that the semiaxis ratio for the intermediate problem is almost constant and close to unity; the ellipse is slightly extended along the Y axis. Numerical estimates obtained using (16) give $D_{x,y} \approx 7.7$ m; i.e., the linear size of the ellipse is $2D_{x,y} \approx 15.4$ m.

High-accuracy trajectory measurements of the position of the pole made in the mid-1990s [7] give trajectories resembling an elliptical (nearly circular) spiral, for which the double angular values of the semiaxes are approximately $0.''45$–$0.''5$, with the corresponding linear values being 14.2–15.7 m. In addition, the elliptical spiral is slightly extended along the Y axis. Therefore, the results of our study and estimates of the Chandler polar motion show quite good qualitative and quantitative agreement with the experimental data, confirming the adequacy of the mathematical model used. It is well known that the average period of the Chandler polar motion is 433 sidereal days, and that this motion represents a free nutation of the viscoelastic Earth, with the corresponding inertia tensor [1–4, 7]. There is also a forced polar motion with a period of one year (about 365 sidereal days), due, in our opinion, to the motion of the Earth–Moon barycenter in its orbit around the Sun. It has a regular character and can be explained by a celestial-mechanical model. The polar oscillations with these periods are usually considered to be the basic ones in the coordinates x_p and y_p of the Earth, derived from observations conducted from the end of the 19th century to the present time. There is also a secular displacement (drift) of the pole along the y_p axis with a velocity of about $0.''033$ per year, which has led to a shift of the oscillatory figure toward North America by approximately 10 m. The corresponding drift along the x_p axis is negligible.

Let us use the astrometric data indicated above to construct a mathematical model of the Earth's polar motion, taking into account three basic factors—the drift, Chandler motion, and annual oscillations. Summing corresponding components along the x_p and y_p axes, we obtain the desired expressions in the principal approximation

$$-x_p = c_x + a_x\cos(n_{w1}t+\Delta_{1x}) \qquad (17)$$

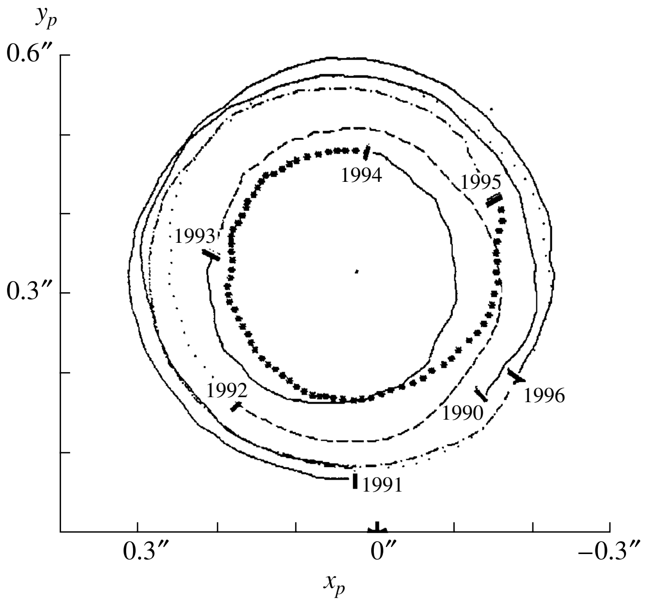

Fig. 3. Trajectory of polar motion from IERS measurements for 1990–1997 taking into account the adopted orientation of the coordinate axes (the x axis is directed toward Greenwich and the y axis 90° to the West).

$$+ b_x \cos(n_h t + \Delta_{hx}),$$
$$y_p = c_y + a_y \sin(n_{w1} t + \Delta_{1y})$$
$$+ b_y \sin(n_h t + \Delta_{hy}),$$

where n_h is the frequency of the annual oscillation, and we have used the directions of the axes adopted by the International Earth Rotation Service (IERS).

The drift $c_{x,y}$, amplitudes $a_{x,y}$, $b_{x,y}$ and phase shifts $\Delta_{1x,y}$, $\Delta_{hx,y}$ can be derived from high-accuracy IERS astrometric data for the end of the 20th century. It is natural to assume that $a_x \approx a_y = a$, $b_x \approx b_y = b$, $\Delta_{1x} \approx \Delta_{1y} = \Delta_1$, and $\Delta_{hx} \approx \Delta_{hy} = \Delta_h$. The numerical values yielded by measurements of the trajectory of the North pole for 1990–1997 referenced to epoch 1900 are approximately (Fig. 3)

$$c_x \approx -0.''02, c_y \approx 0.''33, a \approx 0.''205,$$
$$b \approx 0.''06, \Delta_1 \approx -0.18\pi, \Delta_h \approx -\pi/2.$$

Substituting these values into the polar-motion model (17), we can construct a theoretical trajectory (Fig. 4) corresponding to the observational data. There is a fairly good qualitative agreement between these curves, so reliable forecasts of the polar motion can be performed for about 5 years.

Note that, on average, the experimental and theoretical curves display the same time variations. Of course, the measurements contain perturbing factors with various physical origins, such as measurement and processing errors, as well as geophysical and gravitational perturbations. In particular, the deviation of the experimental curve in the middle of 1994 can be explained by a considerable increase in the influence of Jupiter. A grand alignment of planets can also produce irregular perturbations.

Using the mathematical model (17), we can find the range of oscillatory motions of the pole with respect to the Earth's figure axis and the angular velocity of the precession δn:

$$(a - b)^2 \leq (x_p + c_x)^2 + (y_p - c_y)^2$$
$$\leq (a + b)^2, \delta n = n_h - n_{w1}.$$

These characteristics are somewhat more stable than the trajectory measurements over the time intervals considered.

5. MODELLING THE DISSIPATIVE EVOLUTION OF ROTATIONS OF THE DEFORMABLE EARTH

Let us present a rational approach to the development and analysis of a perturbed dynamical model for the rotational motions of the deformable Earth [1–7]. We will assume that the perturbations are due to deformations of the viscoelastic mantle (i.e., the outer spherical layer, see above) due to the gravitational tides of the Sun and Moon. The basic aim of developing an evolutionary model is to identify parameters of the Chandler pole motion and to interpolate (backward) and extrapolate (forward) the trajectories. It is of interest to consider time intervals available to observation (about 10 to 100 yr), as well as intervals of the order of the precession period ($T_p = 25700$ yr), and even substantially longer.

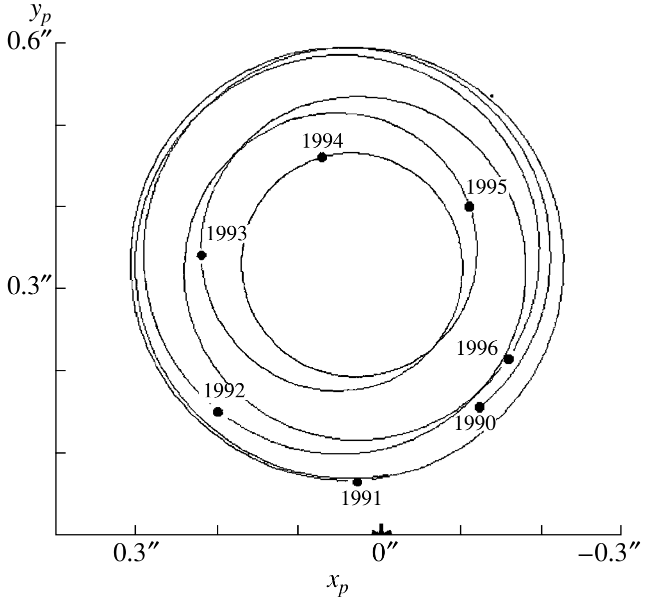

Fig. 4. Theoretical model of the polar trajectory taking into account the drift, Chandler motion, and annual oscillations. The directions of the axes are the same as in Fig. 3.

We introduce the perturbed Routh functional [5, 6]:

$$R = R_0 + \varepsilon R_1(\{I\}, \{w\}, [\mathbf{u}], [\dot{\mathbf{u}}]) + \varepsilon^2 \ldots, \qquad (18)$$

where εR_1 is the perturbed functional associated with the gravitational tides, and $\varepsilon > 0$ is a small dimensionless parameter characterizing the relative amplitude of the perturbing factors in (18), introduced for convenience in the asymptotic analysis. Applying several simplifying assumptions, we can present its structure in the form [5]:

$$\varepsilon R_1 = -\sigma k_0(4\varkappa_1\varkappa_2 + \varkappa_1^2 - \varkappa_2^2) \qquad (19)$$
$$- \sigma k_1(4\gamma_1\gamma_2 + \gamma_1^2 - \gamma_2^2).$$

The coefficients σk_0 and σk_1 in (19) characterize the moments of the gravitational tides due to the Sun and Moon, respectively. They can be expressed as

$$\sigma k_0 = \sigma_* m_0^2 R_s^{-6}, \quad \sigma k_1 = \sigma_* m_1 R_{21}^{-6}. \qquad (20)$$

Here, $R_s = 149.6 \times 10^9$ m is the distance from the Earth to the Sun.

The dimensional parameters σ and σ_* (which are functionals of $\mathbf{u}$ and $\dot{\mathbf{u}}$) must be specified in more detail, and estimated on the basis of astrometric and geophysical measurements. They can be determined only very approximately. It is interesting that $k_0/k_1 \approx 0.2$; i.e., the perturbing influence of the Sun is less than that of the Moon, due to its rapid decrease with distance, in accordance with (20): $R_s \approx 1.49 \times 10^8$ km, $R_{21} \approx 3.84 \times 10^5$ km.

It is of considerable interest to analyze factors in (19) possessing the same structure with respect to the direction cosines $\{\varkappa_j\}$ and $\{\gamma_j\}$ of the basis vectors $\mathbf{R}_s^0$ and $\mathbf{R}_{21}^0$, respectively (see (4) and [5, 6]).

Since the second term is the determining one in (19) and (20), we will consider it in more detail. The direction cosines γ_j of the basis vector $\mathbf{R}_{21}^0$ in the $C_2x_1x_2x_3$ coordinate system fixed to the Earth's core can be expressed in terms of the orbital elements and action–angle variables as follows:

$$\gamma_j = d_1\alpha_j + d_2\beta_j + d_3\nu_j, \qquad j = 1, 2, 3, \qquad (21)$$
$$d_1 = \cos h \cos \vartheta + \sin h \cos i \sin \vartheta,$$
$$d_2 = \cos h \cos i \sin \vartheta - \sin h \sin \vartheta,$$
$$d_3 = -\sin \vartheta \sin i,$$
$$\alpha_j = (s_{1j} \cos w_2 - s_{2j} \sin w_2) \cos w_3$$
$$- (s_{1j} \sin w_2 + s_{2j} \cos w_2) \sin w_3 \cos \delta_1$$
$$+ s_{3j} \sin w_3 \sin \delta_1,$$
$$\beta_j = (s_{1j} \cos w_2 - s_{2j} \sin w_2) \sin w_3$$
$$+ (s_{1j} \sin w_2 + s_{2j} \cos w_2) \cos w_3 \cos \delta_1$$
$$- s_{3j} \cos w_3 \sin \delta_1,$$
$$\nu_j = (s_{1j} \sin w_2 + s_{2j} \cos w_2) \sin \delta_1$$
$$+ s_{3j} \cos \delta_1.$$

Here, $S = (s_{kj})$, where $(k, j = 1, 2, 3)$ is the direction-cosine matrix from the fixed coordinate system $C_2x_1x_2x_3$ to the nonrotating system $C_{12}\xi_1\xi_2\xi_3$ [see (4)]. All angular parameters in (21) have already been defined above. To calculate the corresponding expression in (19), only the components γ_1 and γ_2 are necessary. After averaging over the rapid angular variable w_2, corresponding to the diurnal rotation of the Earth ($n_2 \approx 7 \times 10^{-5}$ s^{-1}), we obtain a formula similar to (19) for s_{31} and s_{32}:

$$\begin{aligned} &\langle 4\gamma_1\gamma_2 + \gamma_1^2 - \gamma_2^2 \rangle_{w_2} \\ &= (4s_{31}s_{32} + s_{31}^2 - s_{32}^2)(d_1 \sin w_3 \sin \delta_1 \\ &\quad - d_2 \cos w_3 \sin \delta_1 + d_3 \cos \delta_1)^2. \end{aligned} \tag{22}$$

It follows from (22) that, at the next stage, we must analyze and calculate the two elements s_{31} and s_{32} of the matrix (s_{kj}), which are determined via the action variables I_1, I_2 and the angle w_1. The corresponding expressions have the form

$$\begin{gathered} s_{31} = \lambda\chi^{-1} \operatorname{cn}(2\mathbf{K}(\lambda)w_1/\pi, \lambda) \\ (\cos\delta_2 = G_{\xi 3}G^{-1} > 0), \\ s_{32} = -\lambda\varkappa_*\chi^{-1} \operatorname{sn}(2\mathbf{K}(\lambda)w_1/\pi, \lambda). \end{gathered} \tag{23}$$

Since $\lambda \ll 1$, accuracy sufficient for our subsequent calculations will be retained if the elliptic functions in (23) are replaced by trigonometric functions or if we use more accurate expansions in λ or $\lambda/\varkappa$. Deriving such expansions with accuracy to $(\lambda/\varkappa)^4$ (i.e., with an error of the order of $(\lambda/\varkappa)^6 \approx 10^{-36}$, we obtain

$$\begin{aligned} &4s_{31}s_{32} + s_{31}^2 - s_{32}^2 \\ &= \left(\frac{\lambda}{\varkappa}\right)^2 (-\varkappa_*^2 + (2 + \varkappa^2)\cos^2 w_1 \\ &\quad -4\varkappa_* \sin w_1 \cos w_1) \\ &+ \left(\frac{\lambda}{\varkappa}\right)^4 \left(1 + \frac{7}{8}\varkappa^2 - \frac{\varkappa^4}{8} + (\frac{\varkappa^4}{8} - \varkappa^2 - 2)\right. \\ &\quad \times \cos^2 w_1 - \frac{\varkappa^8}{8} \sin w_1 \sin 3w_1 \\ &\quad + \frac{\varkappa^2}{8}\cos 4w_1 + 4\varkappa_* \sin w_1 \cos w_1 \\ &\quad \left. - \frac{\varkappa^4}{4}\varkappa_* \sin 4w_1 \right). \end{aligned} \tag{24}$$

As expected, the considered term of the functional εR_1 (19) contains only even harmonics of the angle w_1, corresponding to the phase of the Chandler polar motion. This variable is not considered to be rapid, and we do not average over this phase in studies of the evolution of the system on time intervals of the order of $10^2 T_1$. In accordance with the terminology adopted in asymptotic methods, we will call this the "semislow" phase. The analogous angular variable h (the longitude of ascending node of the lunar orbit in the $C_{12}\xi_1\xi_2$ plane, with a period of revolution of 18.6 yr) is also taken to be semislow. On the other hand, we will consider the variable ϑ, corresponding to the orbital motion of the Moon about the barycenter with a period of about 27.5 sidereal days, to be relatively rapid. Averaging the first factor of (22) over ϑ, we obtain

$$\begin{aligned} &\langle (d_1 \sin w_3 \sin \delta_1 - d_1 \cos w_3 \sin \delta_1 \\ &+ d \cos \delta_1)^2 \rangle_\vartheta = \frac{1}{2}(\cos^2 h + \sin^2 h \cos^2 i) \\ &\times \sin^2 w_3 \sin^2 \delta_1 + \frac{1}{2}(\cos^2 h \cos^2 i + \sin^2 h) \\ &\times \cos^2 w_3 \sin^2 \delta_1 + \frac{1}{2}\sin^2 i \cos^2 \delta_1 \\ &- \frac{1}{4}\sin 2h(1 + \cos^2 i)\sin 2w_3 \sin^2 \delta_1 \\ &+ \frac{1}{4}\cos(h + w_3)\sin 2i \sin 2\delta_1 \\ &\equiv a^2(h, i, \delta_1, w_3). \end{aligned} \tag{25}$$

The angular variables w_3, δ_1, and i in (25) are slow in the context of our analysis of the Chandler polar motion. After averaging over w_2 and ϑ, part of the perturbing functional εR_1 (19) contains periodic factors that depend on the semislow angular variables h and w_1. We should emphasize that a more rigorous averaging over ϑ should be carried out, taking into account the eccentricity of the lunar orbit; however, we limit our treatment here to the use of (25).

Using the averaged functional εR_1 (19)–(25), we can write the required equations for the evolution of the action–angle and other variables. In particular, we obtain

$$\begin{gathered} \dot I_1 = -\varepsilon\frac{\partial R_1}{\partial w_1}, \quad \dot I_2 = 0, \quad \dot I_3 = -\varepsilon\frac{\partial R_1}{\partial w_3}, \\ \dot w_1 = n_1 + \varepsilon\frac{\partial R_1}{\partial I_1}, \quad \dot w_2 = n_2 + \varepsilon\frac{\partial R_1}{\partial I_1}, \\ \dot\delta_2 = \varepsilon(I_2\varkappa_* \sin\delta_2)^{-1}\frac{1 + \varkappa^2 \operatorname{sn}^2(\eta, \lambda)}{\operatorname{dn}(\eta, \lambda)}\frac{\partial R_1}{\partial w_1}, \\ \eta = \frac{2}{\pi}\mathbf{K}(\lambda)w_1. \end{gathered} \tag{26}$$

Recall that the variable δ_2 is the angle between the Earth's figure axis and the angular momemtum vector; its relation to the Andoyer variables L and G is given by (5). The angle δ_1 in the factor in (25) is specified by (5) in terms of the slow variables $G_{\xi 3}$ and G. Equations (26) enable us to analyze variations in I_j, w_1, w_2, and δ_2 on time intervals of the order of $T_2\varepsilon^{-1} \approx 10^3$–$10^4 T_1$, i.e., comparable to the period of precession of the Earth's axis T_p. The absolute value of the angular momemtum of the Earth is $I_2(t) \equiv I_2^0$, whereas $I_1 = I_1^0(1 + O(\varepsilon n_2/n_1))$ (where $n_2/n_1 \approx 430$) undergoes weak quasi-periodic time variations with periods of about $T_h = 1$ yr and $T_1^* = 430$ d; secular variations are absent. Changes in the frequencies $\dot w_j$ due to the tidal gravitational moments are also quite slow. It is interesting that the small angular variable δ_2 possesses a secular variation: its absolute value decreases in the course of time when s_{31} is described by (23). In fact,

$$\delta_2^2 \approx \delta_2^{02} + \varepsilon\mu t, \quad \varepsilon\mu = -2\frac{\sigma}{I_2}k_1\lambda^2 a^2 < 0, \qquad (27)$$

$$\delta_2^0 \approx 0.''2 \approx 10^{-6}.$$

It follows from (27) that the figure axis C^* tends to the angular momemtum vector G.

The system of equations (26) can be used to study the evolution of the Earth–Moon system in the solar gravitational field on time intervals $t \sim T_2/\varepsilon$ sufficient for considerable changes of the slow variables I_3, w_3, and δ_2, as well as shifts of the phases w_1 and w_2. In such an analysis, we can average over w_1 and h, which yields $I_1(t) \approx I_1^0$; i.e., $\lambda \approx \text{const}$. Therefore, the dynamical characteristics of the Earth's intermediate trajectory prove to be quite stable. Apparently, any appreciable changes in this model for the Earth–Moon system occur on time scales $t \gg T_2/\varepsilon$, i.e., $t \gg T_p$.

6. CONCLUSIONS

(1) We have constructed an intermediate trajectory for the rotation of the deformable Earth about the center of mass taking into account the barycentric distance of the Earth–Moon system. This forms a basis for the development of a high-accuracy theory of the Chandler polar oscillations and comparison with astrometric observations.

(2) We have obtained refined analytical expressions for two frequencies, corresponding to the Chandler motion and diurnal rotation of the Earth.

(3) By estimating the elastic deformation of the Earth and variations in its inertia tensor, we have derived numerical values of the period and amplitudes of the polar oscillations, which are in good agreement with observational data. This qualitative and quantitative agreement between the theoretical and experimental data confirms the adequacy of the adopted mathematical model.

(4) Using asymptotic methods of nonlinear mechanics and mathematical modeling of the equations of motion in the action–angle variables, we have carried out a qualitative analysis and derived for the first time quantitative estimates of the complex dynamical process responsible for the mutual locations of the instantaneous axis of the Earth's rotation, its figure axis, and the angular momemtum vector. The dynamical characteristics of the intermediate trajectory of the Earth proved to be quite stable.

ACKNOWLEDGMENTS

This study was supported by the State Science and Technology Program "Astronomy" (project no. 1.8.12) and the Russian Foundation for Basic Research (projects nos. 01-02-17250, 01-01-06306, and 99-01-00276).

REFERENCES

1. H. Moritz and I. I. Mueller, *Earth Rotation: Theory and Observation* (Ungar, New York, 1987; Naukova Dumka, Kiev, 1992).
2. W. Munk and G. Macdonald, *The Rotation of the Earth* (Cambridge Univ. Press, Cambridge, 1960; Mir, Moscow, 1964).
3. Yu. N. Avsyuk, *Tidal Forces and Natural Processes* [in Russian] (Ob"ed. Inst. Fiz. Zemli Ross. Akad. Nauk, Moscow, 1996).
4. G. S. Kurbasova and L. V. Rykhlova, Astron. Zh. **72**, 945 (1995) [Astron. Rep. **39**, 845 (1995)].
5. Yu. G. Markov and I. S. Minyaev, Kosm. Issled. **32** (6), 89 (1994).
6. L. I. Konkina, Yu. G. Markov, A. M. Mikisha, and L. V. Rykhlova, Dokl. Akad. Nauk **375**, 619 (2000) [Dokl. Phys. **45**, 672 (2000)].
7. IERS Annual Reports 1990, July 1991 to 1997, July 1998. Central Bureau of IERS.
8. A. A. Il'yushin, *Mechanics of Continuous Media* [in Russian] (Mosk. Gos. Univ., Moscow, 1990).
9. F. L. Chernous'ko, Izv. Akad. Nauk SSSR, Mekh. Tverd. Tela, No. 1, 22 (1980).

Translated by Yu. Dumin