The Role of a Magnetic Field in the Formation of Jet-like Features in the Crab Nebula

D. V. Khangoulian[1,2*] and S. V. Bogovalov[1]

[1]*Moscow Institute of Engineering Physics (Technical University),*
Kashirskoe shosse 31, Moscow, 115409 Russia
[2]*Max-Planck-Institut fur Kernphysik Saupfercheckweg 1, 69117 Heidelberg, Germany*
Received January 15, 2003

Abstract—The toroidal magnetic field frozen in the relativistic plasma ejected by pulsars must play a significant role in the formation of jet-like features observed in the central parts of plerions. We performed a semiquantitative analysis and calculations of the plasma flow in a plerion using the perturbation theory. We show that for the latitudinal magnetic-field distribution expected during the interaction of the pulsar wind with the interstellar medium, the magnetic field will have an appreciable effect on the flow primarily near the rotation axis. In the equatorial region, the effect of the magnetic field is negligible up to distances of $7r_{\rm sh}$.
© 2003 MAIK "Nauka/Interperiodica".

Key words: *plasma astrophysics, magnetohydrodynamics, and shock waves.*

INTRODUCTION

The Crab Nebula is one of the most interesting and best studied sources in the sky. This object has been observed over a wide wavelength range, from radio to gamma rays with a photon energy of 50 TeV (Aharonian *et al.* 2000; Hester 1998; Shklovskii 1968). One of the most interesting peculiarities of the Crab Nebula is the discovery of a toroidal structure surrounding the pulsar PSR 0531+21 in its central part and two jet-like features arranged perpendicularly to the torus and emerging from the pulsar (Weisskopf *et al.* 2000). Such a structure is not something special. It was also detected in other plerions. Chandra observations revealed similar structures in the Vela (Pavlov *et al.* 2000, 2001; Helfand *et al.* 2001) and PSR 1509−58 (Kaspi *et al.* 2001) plerions and in the supernova remnants G0.9+1 (Gaensler *et al.* 2001) and G54.1+0.3 (Lu *et al.* 2002).

The mechanism of the formation of the Crab Nebula through the ejection of $e^{\pm}$ pairs from a pulsar and the generation of a toroidal magnetic field in a plerion was suggested by Kardashev (1964). These ideas formed the basis of the present-day approach to interpreting the toroidal structure observed in the central part of the Nebula (Aschenbach and Brinkmann 1975). Recently, we have offered (Bogovalov and Khangoulian 2002a) a simple and natural explanation of the formation of toroidal and jet-like structures in plerions based on the MHD model

of Kennel and Coroniti (1984). The new element that was added to the model is a realistic latitudinal distribution of the energy flux density in the pulsar wind. This element made it possible to easily reproduce the toroidal structure of the synchrotron radiation in plerions. In this model, jet-like features are obtained along the rotation axis if the postshock plasma is assumed to spread more or less radially. The surface brightness of the synchrotron radiation is then close to the observed one (Bogovalov and Khangoulian 2002b).

To improve the suggested model, we are currently performing numerical simulations of the interaction of the pulsar wind with the interstellar medium (Bogovalov *et al.* 2002). Our simulations in the hydrodynamic approximation (when the effect of a magnetic field on the plasma dynamics is disregarded) confirm our conclusion about the formation of a toroidal structure. However, the effect of plasma-flow compression to the equatorial plane is so strong that all of the postshock plasma flow transforms into a disklike flow along the equatorial plane. There is absolutely no plasma flow along the rotation axis. This raises the question of how a toroidal magnetic field affects the formation of jets in a plerion. The effect of a magnetic field on the dynamics of a subsonic plasma flow was discussed in several papers in connection with various astrophysical phenomena. In particular, the role of a magnetic field was considered by Bisnovatyi-Kogan (1970) in connection with a magnetorotational supernova explosion. This question was qualitatively

discussed by Lyubarsky (2002) in connection with plerions. Here, our goal is to quantitatively analyze the situation.

ESTIMATING THE EFFECT OF A MAGNETIC FIELD ON THE PLASMA DYNAMICS IN THE CRAB NEBULA

Analysis of the observations of the Crab Nebula indicates that the pulsar wind from PSR 0531+21 is weakly magnetized. The ratio of the magnetic-energy density to the kinetic-energy density of the plasma is only 0.3%. Therefore, in our previous study (Bogovalov and Khangoulian 2002a), we disregarded the effect of a magnetic field on the passage of the pulsar wind through a shock wave and on the plasma flow in a plerion. In this paper, we first give simple quantitative estimates of the magnetic-field effect on the postshock flow and then solve this problem using the perturbation theory.

The magnetic-field distribution in the preshock pulsar wind can be determined from the condition of magnetic-field freezing in plasma $E + \frac{1}{c}[v \times B] = 0$. The electric field in the wind is related to the poloidal magnetic field B_p by $E = \frac{r \sin \theta \Omega}{c} B_p$ (Mestel 1968), where r is the distance to the pulsar, θ is the polar angle, and Ω is the angular velocity of the pulsar. Then,

$$r \sin \theta \Omega B_p + B_\varphi v_p = v_\varphi B_p. \tag{1}$$

At distances $r \gg c/\Omega$, the angular velocity of the plasma tends to zero, because the angular momentum of the plasma particles $(r \sin(\theta) v_\varphi)$ is limited above. In this limit, the toroidal magnetic field at large distances from the pulsar is

$$B_\varphi = \frac{r \sin \theta \Omega}{v_p} B_p. \tag{2}$$

To estimate the preshock magnetic field, we assume that during the motion away from the pulsar, the same fraction of the initial Poynting vector flux transforms into the kinetic energy of the plasma. The ratio of the energy flux density of the electromagnetic field to the kinetic-energy flux density of the plasma will then be independent of the angle θ, except for a narrow region near the rotation axis where the kinetic-energy flux density exceeded the energy flux density of the electromagnetic field from the very outset. To take this circumstance into account, we introduced the additional factor $\frac{\gamma_m \sin^2 \theta}{\gamma_1}$ on the right-hand side of relation (4):

$$\frac{B_a(\theta)^2}{4\pi n_1 m c^2 \gamma_1^2} = \sigma \frac{\gamma_m \sin^2 \theta}{\gamma_1}. \tag{3}$$

Here, $n_1(r)$ is the preshock plasma density, $\gamma_1 = \gamma_0 + \gamma_m \sin^2 \theta$ is the Lorentz factor, and $B_a(\theta)$ is the toroidal magnetic field at distance a from the pulsar. The value of a can be determined from the condition that particle acceleration takes place at $r < a$. Then, the following relation holds:

$$B_a(\theta) = \sqrt{4\pi \sigma m c^2 \gamma_1 n_1 \gamma_m} \sin \theta, \tag{4}$$

with $\gamma_1 n_1$ being independent of the polar angle.

The radius of the shock front $r_{sh}(\theta)$ is defined as (Bogovalov and Khangoulian 2002a)

$$r_{sh} = \sqrt{\frac{3}{2} \left(\frac{3}{\sqrt{8}}\right)^{\frac{\delta}{\delta-1}} \frac{\gamma_1 m c^2 \dot{N}}{4\pi P_{ext} c}}, \tag{5}$$

where P_{ext} is the pressure of the interstellar medium, $\delta = 4/3$ is the polytropic index of an ultrarelativistically hot plasma, and $\dot{N}$ is the rate of particle injection by the pulsar.

In the limit $\gamma_m \gg \gamma_0$ of interest, we can assume that everywhere, except for the narrow range of angles $\theta < \sqrt{\gamma_0/\gamma_m}$, the radius of the shock front is $r_{sh} = r_{eq}|\sin \theta|$. It follows from expression (4) and the freezing-in condition that the preshock magnetic field is

$$B_1 = \mathrm{const} \frac{a}{r_{sh}} \sin \theta. \tag{6}$$

Taking into account the expression that defines r_{sh}, we see that the magnetic field will be constant on the shock surface. At the shock, the magnetic field increases by a factor of 3. As long as the field remains weak, it is easy to understand what will happen to it subsequently. Because of the freezing-in condition, the ratio of the magnetic-field strength to the density will be proportional to the field-line length (Landau and Lifshitz 1982). For a toroidal magnetic field, the field-line length is proportional to the distance from the rotation axis. Since the plasma density varies only slightly in the postshock subsonic flow, the postshock magnetic field on a streamline can be written as

$$B_2 = B_{sh} \left(\frac{\chi}{\chi_{sh}}\right). \tag{7}$$

Here, $\chi_{sh} = r_{sh} \sin \theta$ is the distance to the rotation axis at the shock, χ is the distance to the rotation axis at an arbitrary postshock point, and B_{sh} is the magnetic field immediately behind the shock. It is important to bear in mind that this expression does not depend on the form of the postshock streamlines.

Let us estimate the ratio of the magnetic pressure to the hydrodynamic pressure:

$$\frac{p_m}{p_h} = \frac{1}{8\pi} \frac{B_{\mathrm{sh}}^2 \left(\frac{r}{r_{\mathrm{sh}}}\right)^2}{\frac{2}{3} n_1 m c^2 \gamma_1^2} \qquad (8)$$

$$= \frac{27}{4} \sigma \left(\frac{\gamma_m \sin^2 \theta}{\gamma_1}\right) \left(\frac{\chi}{\chi_{\mathrm{sh}}}\right)^2.$$

For the Crab Nebula, $\sigma = 3 \times 10^{-3}$. Therefore, the magnetic field for the Nebula begins to affect the plasma dynamics when the streamline goes to distance χ from the rotation axis,

$$\chi \sim 7 \chi_{\mathrm{sh}}. \qquad (9)$$

Since $\chi_{\mathrm{sh}} \approx r_{\mathrm{eq}} \sin^2 \theta$, it is clear that for the streamlines near the equator, the magnetic field in the formation region of the toroidal structure in the Crab Nebula is of no importance, because this region lies within $3-4 r_{\mathrm{eq}}$. Recall that the shock radius for the Crab Nebula at the equator is 0.1 pc, while the outer radius of the torus is $\sim$0.4 pc (Weisskopf *et al.* 2000). In contrast, for all of the streamlines near the rotation axis that emerge from the pulsar within the polar angle $\sin \theta < \frac{1}{\sqrt{7}}$, the magnetic field will affect the postshock plasma dynamics even when they lie within $\chi = r_{\mathrm{eq}}$ from the rotation axis.

Partially, the effect of magnetic field on the plasma dynamics becomes clear even from the study of Kennel and Coroniti (1984). When the magnetic pressure becomes equal to the plasma particle pressure, the magnetic field ceases to grow and the dynamics of the plasma motion changes. However, it is obvious that this is not all. Kennel and Coroniti (1984) assumed that the postshock plasma spreads purely radially. Clearly, when the magnetic pressure becomes equal to the plasma-pressure, the radial pattern of motion also breaks down. To understand the nature of this breakdown, it is interesting to consider the analytical solution of the problem in terms of the perturbation theory.

EFFECTS OF A MAGNETIC FIELD ON POSTSHOCK STREAMLINES

In an actual pulsar wind, the particle energy strongly depends on the latitude. The ratio γ_m/γ_0 is of the order of 10^4. However, to estimate the magnetic-field effect, it is interesting to consider the flow for an isotropic kinetic energy flux of particles but in the presence of a weak magnetic field. This can be done in terms of the perturbation theory. In this case, it is important to take into account the fact that for a Lorentz factor $\gamma > 200$, the magnetic collimation

of the preshock plasma is negligible (Beskin *et al.* 1998; Bogovalov and Tsinganos 1999; Bogovalov 2001). The wind from the pulsar may be assumed to spread radially, because at relativistic velocities of the preshock plasma with $\gamma > 200$, the magnetic-field effect (the Ampère force $[j \times B]$) is actually completely offset by the electric field (the Coulomb force). Estimates indicate that the angle through which the streamline deflects from the radial direction is of the order of $\delta\theta = \frac{\sigma}{\gamma_0^2} \ln \frac{r_{\mathrm{sh}}}{R_l}$, where R_l is the radius of the light cylinder (Bogovalov 2001). For the Crab Nebula, $\frac{r_{\mathrm{sh}}}{R_l} \sim 10^9$. Therefore, the deflection of motion from the radial direction upstream the shock is only $\delta\theta = 1.5 \times 10^{-6}$ for $\gamma_0 = 200$. In practice, this implies that the magnetic collimation of the preshock pulsar winds may be completely disregarded.

The situation changes after the shock passage. Qualitatively, this can be understood in the case of a perpendicular shock wave. The collimating force is proportional to the difference $B_{\mathrm{sh}}^2 - E^2$. Upstream the shock, $E = vB_{\mathrm{sh}}/c$. Therefore, the collimating force is proportional to $B_{\mathrm{sh}}^2/\gamma^2$. After the shock passage, the electric field does not change, while the magnetic field in a strong shock almost triples. As a result, the Ampère force immediately behind the shock is almost a factor of 9 larger than the Coulomb force (the forces are proportional to the squares of the fields). The collimating force becomes proportional to $8B_{\mathrm{sh}}^2$. This sharp increase in the collimating force is of considerable interest to us.

The problem is axisymmetric. Let us introduce the stream function ψ. It is related to the physical quantities by

$$n u_r = \frac{1}{r \sin \theta} \frac{\partial \psi}{r \partial \theta}, \qquad (10)$$

$$n u_\theta = -\frac{1}{r \sin \theta} \frac{\partial \psi}{\partial r}, \qquad (11)$$

where n is the proper particle number density and u_r and u_θ are the corresponding four-velocity components. It is convenient to use the equations for the stream function in spherical coordinates. The equation in question is an analogue of the Grad–Shafranov equation (the material on this subject is available in the review by Beskin 1997). It follows from the complete system of magnetohydrodynamic equations for the flow geometry under consideration, assuming that the magnetic field and the deviation of the flow from a spherically symmetric one are small:

$$\frac{u^2}{u_s^2 - u^2} \left\{ \left(u_s^2 - u_r^2\right) \psi_{rr} + \left(u_s^2 - u_\theta^2\right) \frac{\psi_{\theta\theta}}{r^2} \right. \qquad (12)$$

$$-2u_\theta u_r \frac{\psi_{r\theta}}{r} - u_s^2 n u_r \cos\theta$$

$$\left. + n u_\theta \sin\theta \left(2u_r^2 + u_\theta^2\right)\right\} = \frac{u^2 n^2 r^2 \sin^2\theta}{u_s^2 - u^2}$$

$$\times \left[\frac{\delta-1}{\delta}\left(u_s^2 - u^2\right)\frac{\partial \ln S}{\partial \psi} + \gamma^2 u_s^2 \frac{\partial \ln A/S}{\partial \psi}\right]$$

$$+ \frac{u_s^2}{u_s^2 - u^2}\frac{1}{4\pi A(u_s^2 - u^2)}\frac{B^2}{\gamma n}\psi_i u_k \frac{\partial u_k}{\partial x^i}$$

$$+ \frac{\gamma^2 u_s^2}{(u_s^2 - u^2)}\frac{1}{4\pi A}\frac{B^2}{\gamma n}\psi_i$$

$$\times \left(\frac{\partial \ln S}{\partial \psi}\psi_i + 2\frac{1}{\gamma^2}u_k \frac{\partial u_k}{\partial x^i} - \frac{1}{B^2}\mathbf{B}\frac{\partial \mathbf{B}}{\partial x^i}\right)$$

$$- \frac{\psi_i}{w}\left\{\mathbf{E}\rho + \frac{1}{c}\left[\mathbf{jB}\right]\right\}_i.$$

Here, $u_s^2 = \dfrac{\delta-1}{2-\delta}$ and $\delta = 4/3$ is the polytropic index of the ultrarelativistically hot plasma,

$$S(\psi) = \frac{p}{n^\delta}. \tag{13}$$

Since the motion is adiabatic, S is conserved along the streamline and depends only on ψ. The Bernoulli integral

$$A = \frac{w\gamma}{n} + \frac{1}{4\pi}\frac{B^2}{\gamma n} \tag{14}$$

is also conserved along the streamline. Here, w is the thermal function per particle in the intrinsic coordinate system and p is the proper pressure. This form of the Bernoulli integral corresponds to the geometry of the problem under consideration. The complete version of the Grad−Shafranov equation contains five integrals of motion. The absence of three conserved quantities stems from the fact that we consider a region far from the pulsar. Therefore, the effects related to the angular momentum L_z and the corotation angular velocity Ω_F (we use the notation from Beskin 1997) are negligible. The ratio of the particle flux to the magnetic flux η drops out of the equations because the magnetic field is considered as a perturbation.

In the ultrarelativistic case, w and p are related by the equation of state $w = \dfrac{\delta}{\delta-1}p$. In the case of a weakly magnetized wind, we solve Eq. (12) using the perturbation theory.

The equation of the zeroth approximation corresponds to a radial flow in the absence of a magnetic field,

$$u_s^2 \psi_{\theta\theta} - u_s^2 n u_r \cos\theta = 0. \tag{15}$$

Its solution is

$$\psi = u_2 n_2 r_{sh0}^2 (1 - \cos\theta), \tag{16}$$

where u_2 and n_2 are the plasma velocity and density at r_{sh0}, respectively. In the first order of the perturbation theory, the stream function is

$$\psi = u_2 n_2 r_{sh0}^2 (1 - \cos\theta + \sigma_0 f(r,\theta)), \tag{17}$$

where $\sigma_0 = \dfrac{B_{eq}^2}{4\pi n_1 u_1 \gamma_1 m c^2}$. Here, B_{eq} corresponds to the preshock equatorial magnetic field. The equation for the correction $f(r,\theta)$ can be written as

$$\sigma_0 n_2 u_2 r_{sh0}^2 \left\{\left(u_s^2 - u^2\right) f_{rr}\right. \tag{18}$$

$$\left. + u_s^2 (1 - \eta^2)\frac{f_{\eta\eta}}{r^2} - 2u^2 \frac{f_r}{r}\right\} = n^2 r^2 \sin^2\theta$$

$$\times \left[\frac{\delta-1}{\delta}\left(u_s^2 - u^2\right)\frac{\partial \ln S}{\partial \psi} + \gamma^2 u_s^2 \frac{\partial \ln A/S}{\partial \psi}\right]$$

$$+ \frac{\gamma^2 u_s^2}{u^2}\frac{1}{4\pi A}\frac{B^2}{\gamma n}\frac{\partial \ln S}{\partial \psi}\psi_i \psi_i$$

$$+ \frac{u_s^2 - u^2}{w u^2}\frac{1}{2\pi\gamma^2}\frac{\psi_\theta}{r^2}B^2 \cot\theta.$$

The physical quantities on the right-hand side of (18) correspond to an unperturbed flow, $\eta = \cos\theta$, and the magnetic field can be determined from the freezing-in condition (33).

To determine A and S, we use the shock-adiabat relations for a relativistic oblique shock with a non-zero magnetic field. They can be derived from the expressions given by Kennel and Coroniti (1984) using the Lorentz transformations. For a weak magnetic field, the deviations from a radial flow will be small. Therefore, the general expressions can be simplified:

$$u_2^2 = \frac{1}{8}\left(1 + 9\sigma\right), \tag{19}$$

$$\gamma_2^2 = \frac{9}{8}\left(1 + \sigma\right), \tag{20}$$

$$B_2 = 3B_1 \left(1 - 4\sigma\right), \tag{21}$$

$$w_2 = \frac{8}{3}n_1 m c^2 u_1^2 \left(1 - 7\sigma\right), \tag{22}$$

$$n_2 = n_1 u_1 \sqrt{8}\left(1 - \frac{9}{2}\sigma\right), \tag{23}$$

where $\sigma = \dfrac{B_1^2}{4\pi n_1 u_1 \gamma_1 m c^2}$. The physical quantities with the subscript "2" refer to the region immediately behind the shock, and the quantities with the subscript "1" describe the pulsar wind. These relations

correspond to a zero preshock plasma temperature; i.e., $p_1 = 0$, $e_1 = mc^2 n_1$. The thermodynamic quantities of the plasma after the shock passage are defined by relations (19)–(23) and by the assumption that the plasma is injected by the pulsar isotropically; i.e., the wind density is

$$n_1 = \left(\frac{r_{\text{sh}0}}{r}\right)^2 n_w, \tag{24}$$

where n_w is the density at $r_{\text{sh}0}$.

Under the action of a magnetic field, the shock surface can be deformed. Therefore,

$$r_{\text{sh}} = r_{\text{sh}0}\left(1 + \sigma_0 \sum_m R_m P_m(\cos\theta)\right), \tag{25}$$

where P_m are the Legendre polynomials, and R_m are constants. In this case, the preshock plasma density in the first order in σ is

$$n_1 = n_w\left(1 - 2\sigma_0 \sum_m R_m P_m(\cos\theta)\right) \tag{26}$$

and, according to (6), the magnetic field is

$$B_1 = B_{\text{eq}}\frac{r_{\text{sh}0}}{r_{\text{sh}}}\sin\theta = B_{\text{eq}}\frac{r_{\text{sh}0}}{r_{\text{sh}0}} \tag{27}$$

$$\times \left(1 - \sigma_0 \sum_m R_m P_m(\cos\theta)\right)\sin\theta = B_{\text{eq}}\sin\theta.$$

For convenience, a was chosen to be $r_{\text{sh}0}$. Then,

$$u_2{}^2 = \frac{1}{8}\left(1 + 9\sigma_0 \sin^2\theta\right), \tag{28}$$

$$\gamma_2{}^2 = \frac{9}{8}\left(1 + \sigma_0 \sin^2\theta\right), \tag{29}$$

$$B_2 = 3B_1, \tag{30}$$

$$w_2 = \frac{8}{3}n_w mc^2 u_1{}^2 \tag{31}$$

$$\times \left(1 - 7\sigma_0 \sin^2\theta - 2\sigma_0 \sum_m R_m P_m(\cos\theta)\right),$$

$$n_2 = n_w u_1\sqrt{8} \tag{32}$$

$$\times \left(1 - \frac{9}{2}\sigma_0 \sin^2\theta - 2\sigma_0 \sum_m R_m P_m(\cos\theta)\right).$$

The freezing-in condition determines the dependence of the magnetic field downstream of the shock,

$$B = B_2\frac{r}{r_{\text{sh}0}}\frac{n}{n_2} = B_2\frac{r}{r_{\text{sh}0}}\frac{u_2 r_{\text{sh}0}{}^2}{ur^2} \tag{33}$$

$$= 3B_{\text{eq}}\frac{r}{r_{\text{sh}0}}\frac{u_2 r_{\text{sh}0}{}^2}{ur^2}\sin\theta.$$

Expressions (28)–(32) allow us to determine S and the Bernoulli integral downstream of the shock,

$$S = \frac{p_2}{n_2^\delta} = S_0 \tag{34}$$

$$\times \left(1 + \left(\frac{9\delta}{2} - 7\right)\sigma_0 \sin^2\theta + 2(\delta - 1)\sigma_0\right.$$

$$\left.\times \sum_m R_m P_m(\cos\theta)\right).$$

Since the Bernoulli integral is conserved along the streamline,

$$A = \frac{w_1\gamma_1}{n_1} + \frac{1}{4\pi}\frac{B_1{}^2}{\gamma_1 n_1} \tag{35}$$

$$= \frac{w_1\gamma_1}{n_1}\left(1 + \frac{1}{4\pi}\frac{B_1{}^2}{w_1\gamma_1{}^2}\right) = mc^2\gamma_1\left(1 + \sigma_0 \sin^2\theta\right).$$

The expressions on the right-hand side of (18) can be transformed as

$$\frac{u_s{}^2 - u^2}{wu^2}\frac{1}{2\pi\gamma^2}\frac{\psi_\theta}{r^2}B^2 \cot\theta \tag{36}$$

$$= \frac{(u_s{}^2 - u^2)}{\gamma^2}18\sigma_0\frac{w_1 u_1^2 n}{wu}\left(\frac{r}{r_{\text{sh}0}}\right)^2$$

$$\times \left(\frac{u_2 r_{\text{sh}0}^2}{ur^2}\right)^2 \cos\theta \sin^2\theta$$

because the unperturbed flow is radial, $nur^2 = n_2 u_2 r_{\text{sh}0}^2$. Equation (18) then takes the form

$$(u_s^2 - u^2)f_{xx} + u_s^2(1 - \eta^2)\frac{f_{\eta\eta}}{x^2} - 2u^2\frac{f_x}{x} \tag{37}$$

$$= 2(1 - \eta^2)\frac{n^2 x^2}{(nux^2)^2}\left[-\frac{1}{4}(u_s^2 - u^2)\right.$$

$$\times \left(\eta + \frac{1}{3}\sum_m R_m P_m'(\eta)\right)$$

$$\left.+ \gamma^2 u_s^2\left(2\eta + \frac{1}{3}\sum_m R_m P_m'(\eta)\right)\right]$$

$$- \frac{27}{8}\frac{(u_s^2 - u^2)}{\gamma^2}\frac{1}{u^2}\left(\frac{1}{ux^2}\right)^\delta\left(\frac{1}{ux^2}\right)^2 \eta(1 - \eta^2),$$

where $x = r/r_{\text{sh}0}$.

Let us expand f in the series

$$f(x, \eta) = \sum_k f_k(x)Q_k(\eta), \tag{38}$$

where

$$(1 - \eta^2)Q_{m\eta\eta} = -m(m + 1)Q_m, \tag{39}$$

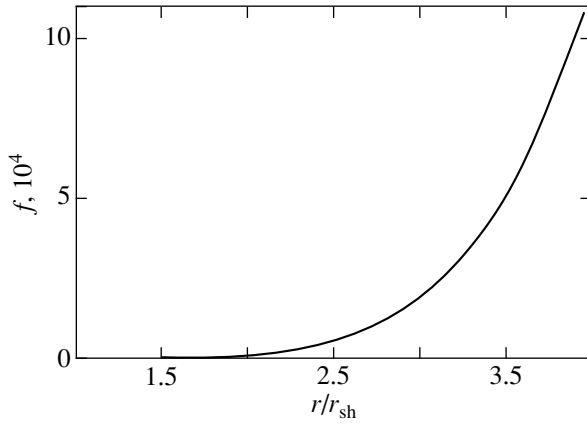$$Q_m = -(1 - \eta^2)P_m'(\eta); \tag{40}$$

Fig. 1. The solution of Eq. (37).

$Q_2 = -3(1 - \eta^2)\eta$ and Eq. (37) takes the form

$$\left(u_s{}^2 - u^2\right) f_{mxx} - u_s{}^2 m(m+1)\frac{f_m}{x^2} - 2u^2\frac{f_{mx}}{x} \tag{41}$$

$$= \frac{\delta_{2m}}{3}\left\{\frac{\left(u_s{}^2 - u^2\right)}{2u^2x^2} - 4\frac{\gamma^2 u_s{}^2}{u^2 x^2}\right.$$

$$\left. + \frac{27}{8}\frac{\left(u_s{}^2 - u^2\right)}{\gamma^2}\frac{1}{u^2}\left(\frac{1}{ux^2}\right)^\delta\left(\frac{1}{ux^2}\right)^2\right\}$$

$$- \frac{2R_m}{3u^2x^2}\left[\gamma^2 u_s{}^2 - \frac{\left(u_s{}^2 - u^2\right)}{4}\right].$$

The general solution of Eq. (41) is

$$f_m(x) = g_m(x) + h(x)\delta_{2m} + R_m j_m(x), \tag{42}$$

where the functions $g_m(x)$, $j_m(x)$, and $h(x)$ satisfy the equations

$$\left(u_s{}^2 - u^2\right) g_{mxx} - u_s{}^2 m(m+1)\frac{g_m}{x^2} - 2u^2\frac{g_{mx}}{x} = 0, \tag{43}$$

$$\left(u_s{}^2 - u^2\right) h_{xx} - 6u_s{}^2\frac{h}{x^2} - 2u^2\frac{h_x}{x} \tag{44}$$

$$= \frac{1}{3}\frac{\left(u_s{}^2 - u^2\right)}{2u^2x^2} - \frac{4}{3}\frac{\gamma^2 u_s{}^2}{u^2 x^2}$$

$$+ \frac{9}{8}\frac{\left(u_s{}^2 - u^2\right)}{\gamma^2}\frac{1}{u^2}\left(\frac{1}{ux^2}\right)^\delta\left(\frac{1}{ux^2}\right)^2,$$

$$\left(u_s{}^2 - u^2\right) j_{mxx} - u_s{}^2 m(m+1)\frac{j_m}{x^2} - 2u^2\frac{j_{mx}}{x} \tag{45}$$

$$= -\frac{2}{3u^2x^2}\left[\gamma^2 u_s{}^2 - \frac{\left(u_s{}^2 - u^2\right)}{4}\right],$$

with the boundary conditions

$$g_m(1) = 0, \tag{46}$$

$$h(1) = 0, \quad h'(1) = 0, \tag{47}$$

$$j_m(1) = 0, \quad j_m{}'(1) = 0. \tag{48}$$

The first boundary condition at the shock front is the continuity of ψ. Therefore,

$$f^{(m)}(1) = 0. \tag{49}$$

The second boundary condition follows from the shock-adiabat relations for an oblique shock (Bogovalov and Khangoulian 2002a):

$$f_m{}'(1) = 2R_m. \tag{50}$$

Then, (42) takes the form

$$f_m(x) = h(x)\delta_{2m} + R_m\left(2j_m(x) + g_m(x)\right) \tag{51}$$

and

$$g_m'(1) = 1. \tag{52}$$

is added to condition (46).

A boundary condition at infinity should be added to these boundary conditions. This additional condition determines the coefficients R_m and, hence, the shape of the shock. However, since the perturbation theory becomes inapplicable at infinity, the solution at $r \gg 1$ is physically meaningless. For a subsonic flow, the perturbations at infinity will have an effect in the entire volume of the flow and on the shape of the shock. However, solution (51) is peculiar in that it is a superposition of the rapidly increasing function $h(x)$, which is nonzero only because the magnetic field is nonzero, and the slowly increasing part $R_2\left(2g_2(x) + j_2(x)\right)$. The function $h(x)$ increases with x as x^6, while the function $R_2\left(2g_2(x) + j_2(x)\right)$ increases only as x^4. Therefore, irrespective of the coefficient R_2, the flow pattern will be determined by the function h. For certainty, we assume that the shock surface is not deformed. Then,

$$R_m = 0 \tag{53}$$

for all values of m. The solution with these boundary conditions is shown in Fig. 1. The behavior of the streamlines that corresponds to this solution is shown in Fig. 2. Thus, a magnetic field causes the streamlines to deviate toward the rotation axis of the pulsar. Recall that we have previously shown (Bogovalov and Khangoulian 2002) that the major factor responsible for the formation of jets is a high plasma density in regions of a small polar angle. Of course, we cannot speak here about the behavior of the solution at large distances from the shock front. Our solution at these distances can yield a qualitatively incorrect result. However, we see that in the jet-formation region at small distances from the shock front, the streamlines deviate toward the polar axis.
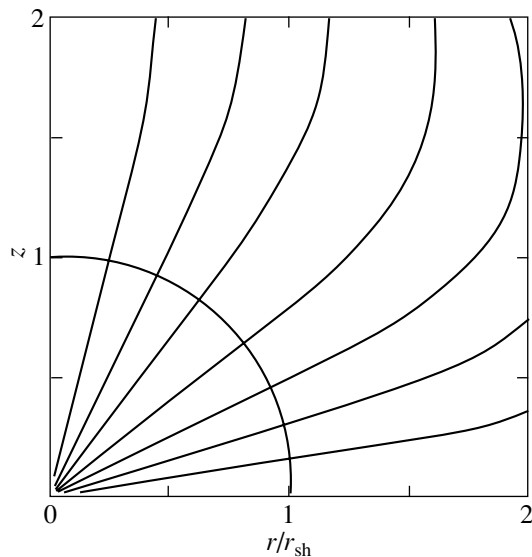
Fig. 2. The distribution of streamlines for a flow with $\sigma = 0.003$. The streamline deviations are an additional factor of the formation of jet-like features.

CONCLUSIONS

We have reached two important conclusions. First, although the preshock energy density of the electromagnetic field is only 0.3% of the kinetic energy density of the plasma, the postshock field increases by a factor of 3 and then grows linearly until the field-energy density becomes equal to the plasma-energy density. Further out, the plasma dynamics changes. Instead of decelerating, the plasma moves at a constant velocity (Kennel and Coroniti 1984). Here, we showed that if we take into account the latitudinal distribution of the energy flux density in the wind from the pulsar, then the magnetic field begins to have an effect primarily on the streamlines located closer to the rotation axis. In this case, its role is probably insignificant in the region of the toroidal structures. Second, the magnetic field changes not only the magnitude of the particle velocity but also the direction of the motion, which results in a collimation of the initially noncollimated plasma flow. Thus, we confirm the idea put forward by Lyubarsky (2002) that the plasma collimation toward the rotation axis in a plerion is one of the factors responsible for the formation of jetlike features in the Crab Nebula. Therefore, the magnetic field should be taken into account in numerical simulations of the interaction of the pulsar wind with the Crab Nebula and the formation of jetlike features.

ACKNOWLEDGMENTS

REFERENCES

1. F. A. Aharonian, A. G. Akhperjanian, J. A. Barrio, *et al.*, Astrophys. J. **539**, 317 (2000).
2. B. Aschenbach and W. Brinkmann, Astron. Astrophys. **41**, 147 (1975).
3. V. S. Beskin, Usp. Fiz. Nauk **167**, 690 (1997).
4. V. S. Beskin, I. V. Kuznetsova, and R. R. Rafikov, Mon. Not. R. Astron. Soc. **299**, 341 (1998).
5. G. S. Bisnovatyi-Kogan, Astron. Zh. **47**, 813 (1970) [Sov. Astron. **14**, 652 (1970)].
6. S. V. Bogovalov and D. V. Khangoulyan, Pis'ma Astron. Zh. **28**, 425 (2002a) [Astron. Lett. **28**, 373 (2002a)].
7. S. V. Bogovalov and D. V. Khangoulian, Mon. Not. R. Astron. Soc. **336**, L53 (2002c).
8. S. V. Bogovalov, A. Koldoba, and G. Ustyugova, in preparation (2002).
9. S. V. Bogovalov and K. Tsinganos, Mon. Not. R. Astron. Soc. **305**, 211 (1999).
10. S. V. Bogovalov, Astron. Astrophys. **371**, 1155 (2001).
11. B. M. Gaensler, M. J. Pivovaroff, and G. P. Garmire, Astrophys. J. Lett. **556**, L107 (2001).
12. D. J. Helfand, E. V. Gotthelf, and J. P. Halpern, Astrophys. J. **556**, 380 (2001).
13. J. J. Hester, in *Neutron Stars and Pulsars: Thirty Years after Discovery,* Ed. by N. Shibazaki *et al.* (Universal Academy, Tokyo, 1998), p. 431.
14. N. S. Kardashev, Astron. Zh. **41**, 807 (1964) [Sov. Astron. **41**, 807 (1964)].
15. V. M. Kaspi, M. J. Pivovaroff, B. M. Gaensler, *et al.*, Am. Astron. Soc. Meeting **197**, 8312 (2001).
16. C. F. Kennel and F. V. Coroniti, Astrophys. J. **283**, 710 (1984).
17. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; translation of the 1st ed., Pergamon Press, Oxford, 1960).
18. F. J. Lu, Q. D. Wang, B. Aschenbach, *et al.*, Astrophys. J. Lett. **568**, L49 (2002).
19. Y. E. Lyubarsky, Mon. Not. R. Astron. Soc. **329**, L34 (2002).
20. L. Mestel, Mon. Not. R. Astron. Soc. **138**, 359 (1968).
21. G. G. Pavlov, D. Sanwal, G. P. Garmire, *et al.*, Astron. Astrophys. **32**, 733 (2000).
22. G. G. Pavlov, V. E. Zavlin, D. Sanwal, *et al.*, Astrophys. J. Lett. **552**, L129 (2001).
23. I. S. Shklovskii, *Supernovae* (Wiley, New York, 1968).
24. M. C. Weisskopf, J. J. Hester, A. F. Tennant, *et al.*, Astrophys. J. Lett. **536**, L81 (2000).

Translated by G. Rudnitskii

On the Mechanism of X-ray Emission from Radio Pulsars

I. F. Malov[*]

Pushchino Radio Astronomy Observatory, Astrospace Center, Lebedev Institute of Physics,
Russian Academy of Sciences, Pushchino, Moscow oblast, 142290 Russia
Received February 10, 2003

Abstract—We discuss the correlations between the luminosities of radio pulsars in various frequency ranges and the magnetic fields on the light cylinder. These correlations suggest that the observed emission is generated in outer layers of the pulsar magnetospheres by the synchrotron mechanism. To calculate the distribution functions of the relativistic particles in the generation region, we use a model of quasi-linear interactions between the waves excited by cyclotron instability and particles of the primary beam and the secondary electron—positron plasma. We derive a formula for calculating the X-ray luminosity L_x of radio pulsars. A strong correlation was found between L_x and the parameter $\dot{P}_{-15}/P^{3.5}$, where P is the neutron-star rotation period, in close agreement with this formula. The latter makes it possible to predict the detection of X-ray emission from more than a hundred (114) known radio pulsars. We show that the Lorentz factors of the secondary particles are small ($\gamma_p = 1.5-8.5$), implying that the magnetic field near the neutron-star surface in these objects is multipolar. It follows from our model that almost all of the millisecond pulsars must emit X-ray synchrotron radiation. This conclusion differs from predictions of other models and can be used to test the theory under consideration. The list of potential X-ray radiators presented here can be used to search for X-ray sources with existing instruments. © *2003 MAIK "Nauka/Interperiodica".*

Key words: *pulsars, neutron stars, and black holes; X-ray sources.*

1. INTRODUCTION

By now, X-ray emission from 41 radio pulsars has been detected by space telescopes (Possenti *et al.* 2002). Many attempts have been made to find correlations between the X-ray luminosities L_x and other parameters of these objects. In particular, a positive correlation was found between L_x and the rate of rotational energy losses $\dot{E}_r = I\Omega\dot{\Omega}$ (Becker and Trümper 1997). Here, Ω is the angular velocity of the neutron star and I is its moment of inertia. Malov and Malov (1995) showed that the integrated radio luminosity of the pulsars L_r also increased with increasing $\dot{E}_r$ ($L_r \propto \dot{E}_r^{1/3} \propto \dot{P}^{1/3}P^{-1}$). Such correlations seem quite natural, because the rotational energy of radio pulsars is believed to be the primary energy source for all types of their emission, although the exact dependence of L on $\dot{E}_r$ is still unknown. Thus, Possenti *et al.* (2002) showed that $L_x \propto \dot{E}^{1.34} \propto \dot{P}^{1.34}P^{-4}$. The dependence $L(\dot{E}_r)$ will probably differ for different samples, and we will not consider it below.

There are two points of view on the localization of the region where the X-ray emission of radio pulsars is generated. One group of authors (e.g., Zhang and Harding 2000) believe that this region is located in the acceleration zone, near the neutron-star surface. Other authors (e.g., Romani and Yadigaroglu 1994) assume that the hard (X-ray and gamma-ray) emission originates in the outer gap, on the periphery of the magnetosphere. In our papers (Malov and Machabeli 1999, 2001, 2002; Malov 2001b), we showed that the optical and X-ray emissions of radio pulsars must be generated in the outer layers of the magnetosphere, near the light cylinder, and there is no need to postulate an outer gap and an additional acceleration of relativistic particles.

Here, we once again discuss arguments for the generation of hard emission at large distances from the surface of a neutron star by the synchrotron mechanism (Section 2) and describe our synchrotron model (Section 3). Subsequently, we consider the possibilities for detecting X-ray emission from a number of known radio pulsars (Section 4) and summarize our main results (Section 5).

2. ARGUMENTS FOR THE SYNCHROTRON MODEL

Shortly after the detection of optical emission from the Crab pulsar, Shklovsky (1970) and Pacini (1971)
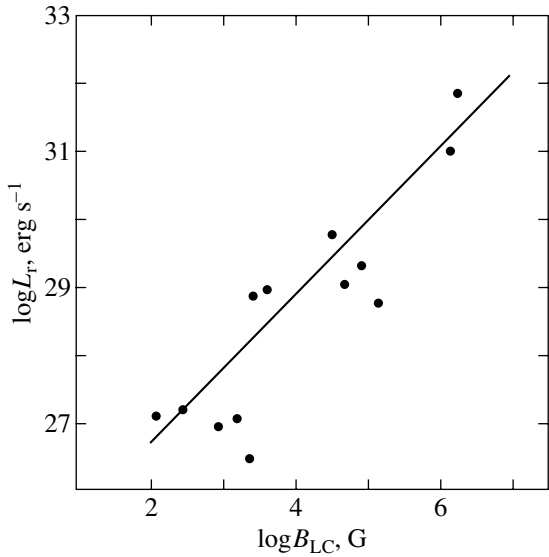
---

[*]E-mail: `malov@prao.psn.ru`

Fig. 1. The relationship between the radio luminosity of pulsars and the magnetic field at the light cylinder.

Fig. 2. The relationship between the optical luminosity of radio pulsars and the magnetic field B_{LC} (the diagram was taken from the paper of Shearer *et al.* 2000).

put forward the idea of using the synchrotron mechanism to explain this emission. They assumed that the generation region was located near the light cylinder ($r \sim R_{LC} = c/\Omega$). Subsequently, a number of authors (Zheleznyakov 1971; Smith 1973; Ferguson 1981) suggested that the radiation from radio pulsars in all wavelength ranges was generated near the light cylinder. This extreme point of view was later transformed into the hypothesis of two types of pulsars (Malov and Suleymanova 1982; Malov 1987). Pulsars with long periods ($P = \Omega/2\pi \sim 1$ s) and with the generation of radio emission inside the magnetosphere at distances $r \ll R_{LC}$ by the curvature radiation mechanism are objects of the first type. Pulsars with short periods ($P \lesssim 0.1$ s) belong to the second group of objects. The observed radiation is generated in these objects near the light cylinder by the synchrotron mechanism. Pulsars with intermediate periods can show properties characteristic of both types. There are various pieces of evidence for this division of radio pulsars (see, e.g., Malov 1997). However, in our view, the most convincing argument for this hypothesis is the high correlation between the integrated radio luminosity L_r and the magnetic field on the light cylinder B_{LC} (Malov 1999) (Fig. 1).

A strong argument for the generation of the entire observed emission in the same region is the similarity of the pulse profiles for the Crab pulsar PSR B0531+21 in all wavelength ranges, from radio to gamma rays (Smith 1977). Another argument for the localization of this region on the light cylinder is the high correlation between the optical luminosity of radio pulsars and the magnetic field B_{LC} (Shearer *et al.* 2000) (Fig. 2).
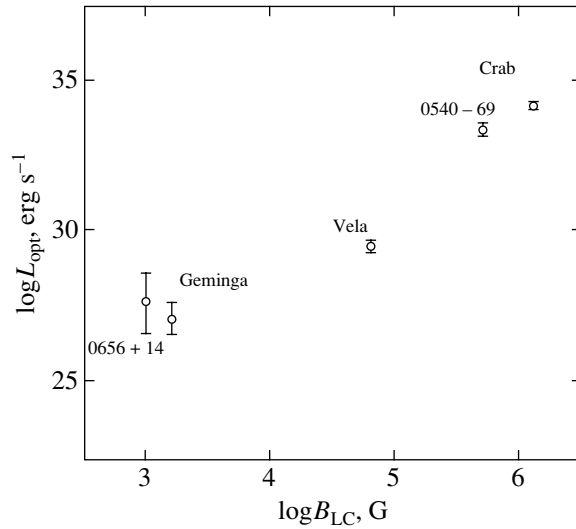
It is interesting to look at the dependence $L_x(B_{LC})$. Using data for 41 pulsars (Possenti *et al.* 2002), we obtain the following relationship between L_x (2–10 keV) and B_{LC} (Fig. 3):

$$\log L_x(\text{erg s}^{-1}) = (1.61 \pm 0.27)\log B_{LC}(\text{G}) \quad (1)$$
$$+ 24.94 \pm 1.25$$

with the correlation coefficient $k = 0.69 \pm 0.12$. This correlation shows that the X-ray emission of radio pulsars is also generated near the light cylinder.

The calculations of the magnetic field at the light cylinder were carried out by assuming a dipole field structure in the entire pulsar magnetosphere:

$$B_{LC} = B_s \left(\frac{R_*}{R_{LC}}\right)^3 = \left(\frac{2\pi R_*}{c}\right)^3 \frac{B_s}{P^3}, \quad (2)$$

where R_* is the neutron-star radius, or for the magnetodipole model of neutron-star braking:

$$B_{LC} = 5.9 \times 10^8 \left(\frac{\dot{P}}{P^5}\right)^{1/2} \text{ G}. \quad (3)$$

It should be emphasized that relation (1) must be corrected with an allowance made for two factors. First, it is necessary to subtract the part associated with the thermal radiation of the neutron-star surface from the observed flux to calculate the nonthermal (synchrotron) luminosity. Second, we should take into account the fact that the angles β between the rotation axes and the magnetic moments $\boldsymbol{\mu}$ of pulsars can be different. In this case, it is necessary to use $B = B_{LC}\sin^3\beta$ (Fig. 4) instead of B_{LC}. It is hoped that an allowance for these two factors will lead to a
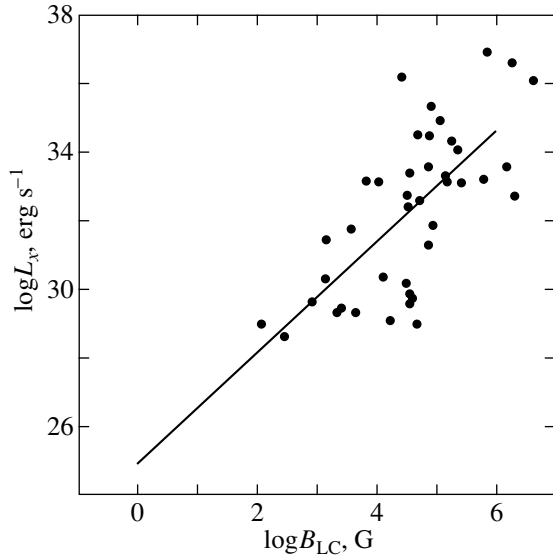
Fig. 3. The observed relationship between the X-ray luminosity of radio pulsars and the magnetic field at the light cylinder.

Fig. 4. The scheme to explain the decrease of the magnetic field strength in the generation region.

higher correlation between $\log L_x$ and $\log B_{LC}$. Note that five radio pulsars with detected optical signals are simultaneously X-ray sources. This implies that their optical radiation is related to the tail of higher-energy emission, and the correlations $L_x(B_{LC})$ and $L_{opt}(B_{LC})$ may not be independent.

The relationships shown in Figs. 1−3 allow us to use the synchrotron model to explain the observational data. This model is described in the next section.

3. DESCRIPTION OF THE MODEL

The total power of the synchrotron radiation emitted by a single electron is given by the expression (Pacholczyk 1970)

$$p = \frac{2e^4 B^2 \gamma^2 \sin^2 \psi}{3m^2 c^3}. \qquad (4)$$

Therefore, the synchrotron luminosity of any object depends on the distribution of the emitting particles in energy ϵ (or Lorentz factor $\gamma = \epsilon/mc^2$), on their pitch angles ψ (the angle between the particle velocity and the magnetic field), and on the magnetic field strength in the region where the radiation is generated. The distribution function shown in Fig. 5 is formed near the neutron-star surface (Arons 1981).

Here, γ_b characterizes the primary beam, and γ_p and γ_t characterize the secondary plasma. This function is one-dimensional, because any transverse momentum is lost by synchrotron radiation in less than 10^{-15} s. However, an anisotropic plasma is unstable
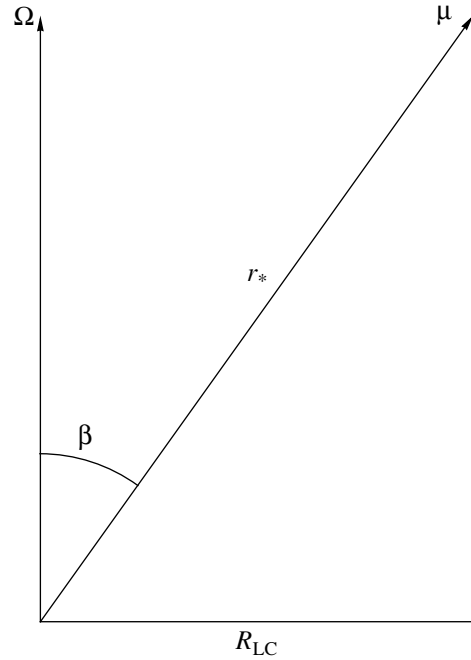
(Sagdeev and Shafranov 1960) and certain types of waves will be excited. Their interaction with particles can give rise to appreciable pitch angles of the particles at sufficiently large distances from the neutron-star surface. Malov and Machabeli (2002) analyzed the steady-state kinetic equation

$$\frac{\partial}{\partial p_{||}} \left\{ \left[\alpha_s \psi_0^2 \left(\frac{p_{||}}{mc} \right)^2 + \alpha_c \left(\frac{p_{||}}{mc} \right)^4 \right. \right. \qquad (5)$$

$$\left. \left. - 2\pi^2 \psi_0 \frac{mc}{p_{||}} r_e \omega_k n_k \right] f_{||} \right\} = 0,$$

where $\alpha_s = 2e^2 \omega_B^2/(3c^2)$, $\alpha_c = 2e^2/(3\rho^2)$, ρ is the radius of curvature of the magnetic field lines, $\omega_B = eB/(mc)$ is the cyclotron frequency, n_k is the number of plasmons, and $r_e = e^2/(mc^2)$ is the electron radius. They took into account the quasi-linear diffusion and the braking forces produced by the synchrotron and curvature radiations.

The solution of Eq. (5) can be sought in the form

$$f(\mathbf{p}) = \chi(\psi) f(p).$$

If the cyclotron resonance is associated with the beam particles, then the distribution functions are

$$\chi(\psi) = C_1 e^{-A_1 \psi^2}, \quad f_{||} \propto \gamma^{-4}. \qquad (6)$$

If the particles of the high-energy tail of the secondary-plasma distribution play a leading role, then the corresponding distribution functions at large distances from the neutron-star surface will be
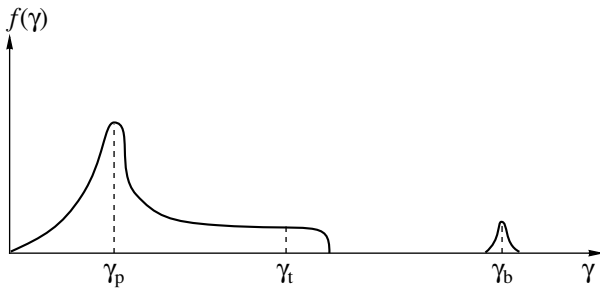
$$\chi(\psi) = C_2 e^{-A_2 \psi^4}, \qquad (7)$$

Fig. 5. The distribution function of relativistic particles in the pulsar magnetosphere.

$$f_{\parallel}^{(1)} \propto \gamma, \quad f_{\parallel}^{(2)} \propto \gamma^{-2}. \tag{8}$$

In formulas (6) and (7), A_1, A_2, C_1, and C_2 are constants. These distributions make it possible to calculate the synchrotron spectra. In particular, if $\gamma_b \sim 10^6$, $\gamma_p \sim 10$, and $\gamma_t \sim 10^5$, then the maximum frequency for PSR B0656+14 should be 7.5×10^{16} Hz (~ 0.5 keV) (Malov and Machabeli 2002), which is consistent with observations (Koptsevich *et al.* 2001). According to dependence (8), the intensity slowly decreases at high frequencies as $\nu^{-0.5}$.

Malov and Machabeli (2002) gave a formula for the synchrotron luminosity of radio pulsars:

$$L = \frac{\pi^2 e^4 I \gamma_r \dot{P} \psi^2 B^2}{4 m^3 c^5 P^2}, \tag{9}$$

where γ_r is the Lorenz factor of the resonant particles. Here, we assume that the synchrotron radiation emanates from the part of the torus that is on the light cylinder and that one-half of the rotational energy losses is transferred by emitting particles. The high-energy tail particles ($\gamma_r = \gamma_t$) mainly contribute to the luminosity, and much of the radiation is emitted in the X-ray range. In this case, the mean Lorentz factor can be estimated from distribution (7); for

$$A_2 = \frac{4 e^6}{3 \pi^3 m^5 c^7} \frac{B^4 P^3 \gamma_p^4 \gamma_r^2}{\gamma_b^3}$$

it is equal to (Malov and Machabeli 2002)

$$\psi_0 \approx \frac{1}{2} \left(\frac{3 \pi^3 m^5 c^7 \gamma_b^3}{4 e^6 B^4 P^3 \gamma_p^4 \gamma_r^2} \right)^{1/4}. \tag{10}$$

The total luminosity can be calculated using the formula

$$L = \frac{\sqrt{3}}{32} \frac{\pi^{7/2} e I \dot{P} \gamma_b^{3/2}}{m^{1/2} c^{3/2} P^{7/2} \gamma_p^2}. \tag{11}$$

The first detailed model of synchrotron radiation from a radio pulsar was developed by Zheleznyakov and Shaposhnikov (1972) for PSR B0531+21. This
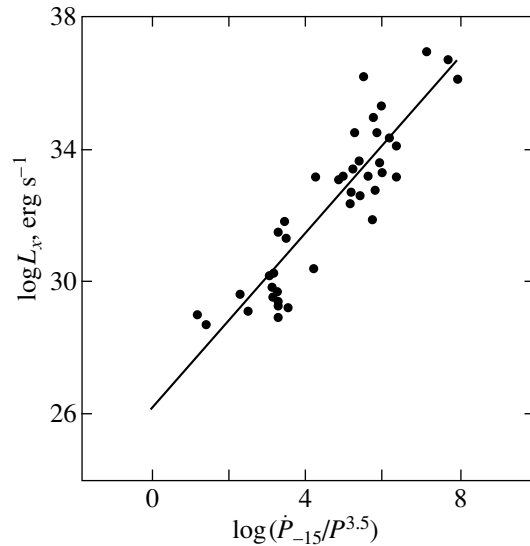
Fig. 6. The observed dependence of the X-ray luminosity of radio pulsars on the parameter $\dot{P}/P^{3.5}$.

model was critically analyzed in our other papers (see, e.g., Malov and Machabeli 2001). Here, we only emphasize that these and other investigators of the synchrotron model arbitrarily assumed the pitch angle to be $\psi \sim 1$. However, $\psi \ll 1$ in the magnetospheres of pulsars (including PSR B0531+21) (Machabeli and Usov 1989; Malov and Machabeli 2001). Indeed, if $\gamma_b = 10^6$, $\gamma_p = 10$, $\gamma_t = 10^5$, $P = 0.1$ s, and $r = R_{LC}$, then $\psi_0 \sim 10^{-3}$.

Formula (11) makes it possible to relate the measured X-ray luminosity of radio pulsars to the observed parameter $\dot{P}/P^{3.5}$ if we assume that $\gamma_b^{1.5}/\gamma_p^2 =$ const. A comparison of L_x and $\dot{P}/P^{3.5}$ for 41 radio pulsars gives (Fig. 6)

$$\log L_x = (1.32 \pm 0.10) \log \frac{\dot{P}_{-15}}{P^{7/2}} + 26.12 \pm 0.48 \tag{12}$$

with the correlation coefficient $k = 0.91 \pm 0.07$.

Dependence (11) describes the relationship between the X-ray luminosity and other parameters of radio pulsars more accurately than the correlation $L_x - B_{LC}$ does. This relationship includes no magnetic field. This is a fortunate circumstance, because the magnetic field on the neutron-star surface B_s is estimated in the magnetodipole model, but it may be incorrect when there are other pulsar slowdown mechanisms (Malov 2001a, 2003). Our subsequent calculations are free from such an uncertainty. In addition, it is not necessary to take into account any difference in the angles β for different pulsars in our model.

The high correlation coefficient between L_x and $\dot{P}/P^{3.5}$ implies that the model under consideration is correct and that the contribution of thermal radiation to L_x is basically small.

4. APPLICATIONS AND PREDICTIONS

The coefficient at $\log \dot{P}/P^{3.5}$ in Eq. (12) slightly differs from unity. If we assume that it is exactly equal to one, then we can calculate the mean value of the parameter $\gamma_b^{1.5}/\gamma_p^2$ for 41 pulsars from relation (11):

$$\log L_x = \log A + \log \frac{\dot{P}_{-15}}{P^{7/2}}. \tag{13}$$

The most probable value of $\log A = 27.6$ corresponds to the parameter

$$\frac{\gamma_{\rm b}^{3/2}}{\gamma_{\rm p}^2} = 4.37 \times 10^8. \tag{14}$$

We obtain $\gamma_p = 1.5-8.5$ for $\gamma_b = 10^6-10^7$. Such Lorentz factors of the secondary plasma can be achieved if the magnetic field near the neutron-star surface has a nondipolar structure (Machabeli and Usov 1989). Indeed, if the radius of curvature of the magnetic field lines is small ($\rho \sim R_*$), then a large number of curvature and synchrotron photons are emitted by each primary electron and much more electron−positron pairs than in the case of a dipole magnetic field, for which $\rho \sim 10^8$ cm $\gg R_*$, are produced. Here, R_* is the neutron-star radius. Therefore, the mean energy (Lorentz factor) of the secondary particles in a multipole field ($\gamma_p \lesssim 10$) is lower than that for a dipole field ($\gamma_p \sim 10^3$).

The minimum X-ray flux F in the range $2-10$ keV is 8×10^{-16} erg s^{-1} cm^{-2} for the sample of 41 pulsars under consideration. Using this value, we can predict the detection of X-ray emission from other radio pulsars in the catalogues of Taylor *et al.* (1995), Manchester *et al.* (2001), and Morris *et al.* (2002). We obtain from formulas (11) and (14)

$$F = \frac{L}{4\pi d^2} = \frac{\sqrt{3}\pi^{5/2} e I \dot{P} \gamma_{\rm b}^{3/2}}{128 m^{1/2} c^{3/2} P^{7/2} \gamma_{\rm p}^2 d^2} \tag{15}$$

$$= 3.34 \times 10^{-17} \frac{\dot{P}_{-15}}{P^{7/2} d_{\rm kpc}^2},$$

which corresponds to the condition

$$\eta = \frac{\dot{P}_{-15}}{P^{7/2} d_{\rm kpc}^2} \geq 24. \tag{16}$$

The table lists the pulsars for which condition (16) is satisfied. We see that one might expect X-ray emission to be detected from more than 100 objects with the current sensitivity. For completeness, we also included pulsars with measured X-ray fluxes (Possenti *et al.* 2002) in the table. They are marked by asterisks. The expected fluxes F were also calculated for them using formula (15). This formula makes it possible to predict the detection of X-ray emission from newly discovered radio pulsars using their periods, the derivatives of the periods, and the distances.

5. DISCUSSION AND CONCLUSIONS

The table shows that almost all of the known millisecond radio pulsars must be X-ray emitters. The very low values of $\dot{P}$ for these objects are offset by the smallness of the factor $P^{3.5}$. Zhang and Harding (2000) concluded that the pulsed X-ray emission from most of the millisecond pulsars had a thermal nature. Future observations of these sources will allow the choice of a correct model to be made.

Below, we summarize our main results.

(1) We gave several arguments for the synchrotron model of X-ray emission from radio pulsars. A correlation was found between the X-ray luminosity and the magnetic field on the light cylinder, suggesting the generation of X-ray emission from these objects near the light cylinder.

(2) We briefly described a model for the formation of the electron-distribution function. This model is based on a quasi-linear interaction of the waves excited by the cyclotron instability in the magnetosphere with particles of the primary beam and the tail of the secondary plasma. A formula was derived to calculate the X-ray luminosity L_x of radio pulsars.

(3) We found a high correlation between L_x and the parameter $\dot{P}_{-15}/P^{3.5}$ (the correlation coefficient is $k = 0.91 \pm 0.07$). This correlation makes it possible to predict the detection of an X-ray flux from known and newly discovered radio pulsars.

(4) We gave a list of 114 objects whose X-ray emission can be detected at the current instrument sensitivity. If the search is successful, the number of radio pulsars with measured X-ray fluxes will at least double.

(5) We showed that the Lorentz factors at the maximum of the secondary-plasma distribution function lie within a very narrow range, $\gamma_p = 1.5-8.5$. This result implies that the magnetic field near the neutron-star surface is multipolar.

(6) The suggested model can be tested by searching for the predicted X-ray fluxes from the radio pulsars listed in the paper. In particular, synchrotron X-ray radiation must be detected from most of the known millisecond pulsars. This prediction is a crucial test that makes it possible to discriminate between the model under consideration and the model

Table

No.	PSR	$\log \eta$	$-\log F$	No.	PSR	$\log \eta$	$-\log F$
1	J0024−0534	4.38	12.10	41*	J1119−6127	<3.57	<12.91
2*	J0030+0451	4.37	12.11	42*	J1124−5916	4.55	11.93
3	B0053+47	1.69	14.79	43	B1133+16	1.45	15.03
4*	B0114+58	<3.59	<12.89	44	B1133−55	1.62	14.86
5	B0136+57	2.09	14.39	45	J1138−6207	1.57	14.91
6*	J0205+6449	5.59	10.89	46	B1143−60	1.39	15.09
7*	J0218+4232	3.56	12.92	47	B1257+12	8.19	8.29
8*	B0355+54	2.84	13.64	48	B1259−63	3.65	12.83
9*	J0437−4715	2.52	13.96	49	J1301−6305	2.59	13.89
10	B0450+55	2.21	14.27	50	B1317−53	1.65	14.83
11*	B0531+21	7.21	9.27	51	B1336−64	1.61	14.87
12*	J0537−6910	4.63	11.85	52	B1338−62	3.03	13.45
13*	J0538+2817	3.02	13.46	53	B1356−60	2.39	14.09
14*	B0540−69	3.85	12.63	54	J1406−6121	2.17	14.31
15	B0540+23	2.22	14.26	55	J1412−6145	1.81	14.67
16	J0613−0200	3.20	13.28	56*	J1420−6048	5.40	11.08
17	B0611+22	2.14	14.34	57*	J1435−6100	1.43	15.05
18	J0631+1036	2.28	14.20	58	B1449−64	2.52	13.96
19*	B0633+17	4.85	11.63	59	B1508−57	1.78	14.70
20*	B0656+14	3.44	13.04	60*	B1509−58	4.77	11.71
21	J0729−1448	2.88	13.60	61	B1516+02A	1.98	14.50
22	B0740−28	3.39	13.09	62	J1530−5327	2.26	14.22
23*	J0751+1807	2.90	13.58	63	B1534+12	2.70	13.78
24*	B0823+26	2.04	14.44	64	B1535−56	1.63	14.85
25*	B0833−45	6.38	10.10	65	J1548−5607	2.03	14.45
26	J0901−4624	1.43	15.05	66	B1556−44	1.65	14.83
27	B0905−51	1.51	14.97	67	B1557−50	1.61	14.87
28	B0906−17	1.61	14.87	68	J1601−5335	2.48	14.00
29	B0906−49	2.94	13.54	69	B1607−52	2.56	13.92
30	B0919+06	<1.47	<15.01	70	B1610−50	3.19	13.29
31	J0940−5428	3.95	12.53	71*	J1617−5055	4.88	11.60
32*	B0950+08	3.29	13.19	72	B1620−26	3.24	13.24
33*	J1012+5307	3.71	12.77	73	B1634−45	2.57	13.91
34*	J1016−5819	2.22	14.26	74	J1640+2224	3.07	13.41
35*	J1024−0719	3.39	13.09	75	J1643−1224	<2.32	<14.16
36	J1045−4509	1.70	14.78	76	B1643−43	2.60	13.88
37*	B1046−58	4.21	12.27	77	B1702−19	2.31	14.17
38*	B1055−52	2.87	13.61	78*	B1706−44	4.92	11.56
39*	J1105−6107	3.71	12.77	79	J1713+0747	3.04	13.44
40	J1112−6103	<2.70	<13.78	80	B1718−35	1.73	14.75

Table (Contd.)

No.	PSR	$\log \eta$	$-\log F$	No.	PSR	$\log \eta$	$-\log F$
81	J1718−3825	3.25	13.23	119	J1839−0321	1.60	14.88
82	B1719−37	2.43	14.05	120	J1841−0348	2.93	13.55
83	J1723−3659	2.06	14.42	121	B1841−05	1.48	15.00
84	B1727−33	3.67	12.81	122	B1842−04	2.00	14.48
85	J1730−2304	3.18	13.30	123	B1844−04	1.50	14.98
86	B1730−37	1.75	14.73	124*	J1846−0258	3.01	13.47
87	B1734−35	1.48	15.00	125	J1849−0317	1.45	15.03
88	B1736−29	1.60	14.88	126	J1853+0056	2.10	14.38
89	B1737−30	2.39	14.09	127*	B1853+01	3.29	13.19
90	J1737−3137	1.83	14.65	128	B1855+09	3.29	13.19
91	J1738−2955	1.98	14.50	129	J1900+0227	2.00	14.48
92	J1739−3023	3.29	13.19	130	J1908+0734	2.75	13.73
93	B1742−30	1.92	14.56	131	J1909+0912	2.01	14.47
94	J1743−3153	1.72	14.76	132	J1913+0832	1.93	14.55
95*	J1744−1134	4.11	12.37	133	J1913+1011	4.27	12.21
96	B1749−28	1.41	15.07	134	B1914+09	1.47	15.01
97	B1754−24	2.23	14.25	135	B1915+13	2.12	14.36
98*	B1757−24	3.87	12.61	136	B1916+14	1.69	14.79
99*	B1800−21	3.99	12.49	137	J1918+1541	2.25	14.23
100	B1802−07	1.43	15.05	138*	B1929+10	3.86	12.62
101	J1809−1917	4.06	12.42	139	B1930+22	2.72	13.76
102*	J1811−1926	4.02	12.46	140*	B1937+21	4.73	11.75
103	B1820−30A	3.65	12.83	141*	B1951+32	4.88	11.60
104	B1820−31	1.94	14.54	142	B1953+29	1.74	14.74
105*	B1821−19	1.82	14.66	143	B1957+20	4.63	11.85
106	B1821−24	4.53	11.95	144	J2010+2425	3.36	13.12
107	B1822−09	2.11	14.37	145	B2020+28	1.68	14.80
108	B1822−14	1.83	14.65	146	B2022+50	1.39	15.09
109*	B1823−13	4.13	12.35	147	J2043+2740	3.59	12.89
110	B1828−10	2.03	14.45	148	B2127+11E	2.41	14.07
111	J1828−1101	3.45	13.03	149	B2127+11F	1.89	14.59
112	B1830−08	3.20	13.28	150*	J2124−3358	4.39	12.09
113	B1832−06	1.80	14.68	151	J2229+2643	2.81	13.67
114	J1835−1020	1.76	14.72	152*	J2229+6114	5.44	11.04
115	J1837−0559	1.56	14.92	153	J2317+1439	2.45	14.03
116	J1837−0604	3.63	12.85	154	J2322+2057	3.32	13.16
117	B1838−04	1.94	14.54	155*	B2334+61	2.57	13.91
118	J1838−0453	1.70	14.78				

* Pulsars with measured X-ray fluxes (Possenti *et al.* 2002).

for the generation of X-ray emission from these objects by the thermal mechanism near the neutron-star surface.

REFERENCES

1. J. Arons, in *Proceedings of the International Summer School and Workshop on Plasma Physics,* Ed. by T. D. Guyenne (Eur. Space Agency, Paris, 1981), p. 273.
2. W. Becker and J. Trümper, Astron. Astrophys. **326**, 682 (1997).
3. D. C. Ferguson, Comments Astrophys. **9**, 127 (1981).
4. A. B. Koptsevich, G. G. Pavlov, S. V. Zharikov, *et al.*, Astron. Astrophys. **370**, 1004 (2001).
5. G. Z. Machabeli and V. V. Usov, Pis'ma Astron. Zh. **15**, 910 (1989) [Sov. Astron. Lett. **15**, 393 (1989)].
6. I. F. Malov, Aust. J. Phys. **40**, 731 (1987).
7. I. F. Malov, Astron. Zh. **74**, 697 (1997) [Astron. Rep. **41**, 617 (1997)].
8. I. F. Malov, Astron. Zh. **76**, 825 (1999) [Astron. Rep. **43**, 727 (1999)].
9. I. F. Malov, Astron. Zh. **78**, 452 (2001a) [Astron. Rep. **45**, 389 (2001a)].
10. I. F. Malov, Astron. Zh. **78**, 990 (2001b) [Astron. Rep. **45**, 865 (2001b)].
11. I. F. Malov, Astron. Zh. (2003), in press.
12. I. F. Malov and O. I. Malov, Astron. Zh. **72**, 574 (1995) [Astron. Rep. **39**, 510 (1995)].
13. I. F. Malov and G. Z. Machabeli, Astron. Zh. **76**, 788 (1999) [Astron. Rep. **43**, 691 (1999)].
14. I. F. Malov and G. Z. Machabeli, Astrophys. J. **554**, 587 (2001).
15. I. F. Malov and G. Z. Machabeli, Astron. Zh. **79**, 755 (2002) [Astron. Rep. **46**, 684 (2002)].
16. I. F. Malov and S. A. Suleymanova, Astrofizika **18**, 65 (1982).
17. R. N. Manchester, A. G. Lyne, F. Camilo, *et al.*, Mon. Not. R. Astron. Soc. **328**, 17 (2001).
18. D. J. Morris, G. Hobbs, A. G. Lyne, *et al.*, astro-ph/0204238 (2002).
19. A. G. Pacholczyk, *Radio Astrophysics* (Freeman, San Francisco, 1970; Mir, Moscow, 1973).
20. F. Pacini, Astrophys. J. Lett. **163**, L17 (1971).
21. A. Possenti, R. Cerutti, M. Colpi, and S. Mereghetti, astro-ph/0109452 (2002).
22. R. W. Romani and I.-A. Jadigaroglu, Astrophys. J. **438**, 314 (1994).
23. R. S. Sagdeev and V. D. Shafranov, Zh. Éksp. Teor. Fiz. **39**, 181 (1960) [Sov. Phys. JETP **12**, 130 (1960)].
24. F. J. Smith, Nature **243**, 207 (1973).
25. F. G. Smith, *Pulsars* (Cambridge Univ. Press, Cambridge, UK, 1977; Mir, Moscow, 1979).
26. A. Shearer, A. Golden, and J. Beskin, in *IAU Colloq. No. 177: Pulsar Astronomy—2000 and Beyond,* Ed. by M. Kramer, N. Wex, and R. Wielebinski (Kluwer Acad., Dordrecht, 2000), p. 307.
27. I. S. Shklovsky, Astrophys. J. Lett. **159**, L77 (1970).
28. J. H. Taylor, R. N. Manchester, A. J. Lyne, and F. Camilo, unpublished work (1995).
29. B. Zhang and A. K. Harding, Astrophys. J. **532**, 1150 (2000).
30. V. V. Zheleznyakov, Astrophys. Space Sci. **13**, 27 (1971).
31. V. V. Zheleznyakov and V. E. Shaposhnikov, Astrophys. Space Sci. **18**, 141 (1972).

Translated by I. Malov

Fission and the r-Process: The Rates of Induced and Delayed Fission

I. V. Panov[1*] and F.-K. Thielemann[2]

[1]Institute for Theoretical and Experimental Physics,
ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia
[2]University of Basel, Basel, Switzerland
Received March 20, 2003

Abstract—The most recent fission-barrier calculations based on improved mass formulas indicate that the adopted values are underestimated. We analyze the dependence of the fission rates on the fission barrier and show that an increase in the fission barriers leads not so much to a decrease in the importance of fission as to the possible synthesis of heavier elements in the r-process. The rates of induced fission of most isotopes with $Z > 80$ at astrophysical energies have been calculated for the first time for fission barriers obtained from different theoretical models. © *2003 MAIK "Nauka/Interperiodica"*.

Key words: *nuclear astrophysics, nucleosynthesis, r-process; supernovae and supernova remnants.*

INTRODUCTION

One of the most important questions in r-process studies is the nucleosynthesis pattern in the region of transuranium nuclei, where fission hampers the passage of the nucleosynthesis wave in the region of actinides and the production of superheavy elements. In particular, an allowance for fission leads to a change in the actinide yields and, hence, to a change in the age of the Galaxy estimated by the isotopic-ratio method.

After the observations of elemental abundances in very old metal-poor stars were published (see, e.g., Sneden *et al.* 2000), it became clear that, at least in one of the naturally occurring r-process scenarios, the characteristic duration of rapid nucleosynthesis τ_r is longer than the time τ_f it takes for the nucleosynthesis wave to reach the region of fissionable nuclei, and because of the high fission rates, the r-process returns to the region of fission-product nuclei with the establishment of a quasi-steady current of nuclei at $\tau_r - \tau_f > \tau_c$ (where τ_c is the cycle time it takes to double the absolute densities of heavy nuclei). More reliable calculations of the fission rates for most transuranium nuclei are required to model this nucleosynthesis.

Analysis of the observations of old metal-poor stars showed that the region of nuclei with masses $110 < A < 130$ is of importance in understanding the role of fission in the r-process. It was noticed that the abundances of these nuclei are much lower than the

r-element abundances in the Solar system. Therefore, the r-element abundance curve must have at least two components: one component dominates at $A > 130$ and the other (weak) component dominates at $A < 130$ (Wasserburg *et al.* 1996). Observations of the main component in metal-poor stars and observations for $A < 130$ can provide appropriate information for determining the fission parameters, because the following explanation of the abundances of elements heavier than iron seems plausible: (i) the expansion of highly neutronized low-entropy matter, which results in the nucleosynthesis looping due to fission and in the formation of nuclei with masses $A \geq 130$ in the main r-process; (ii) the formation of lighter elements in the weak r-process, which proceeds in astrophysical objects of a different type (for more detail, see, e.g., Panov and Chechetkin 2002). However, the underestimated yield of nuclei with masses $A = 100-120$ can also be explained by the peculiarities of the fission-product mass distribution during the fission of nuclei of the dominant r-process component.

Since our prime objective is to determine the contribution of fission to rapid nucleosynthesis, we will not discuss here the possible explanations of the observed heavy-element abundances based on the superposition of computational data (Hill *et al.* 2002; Qian and Wasserburg 2002). We chose the group of scenarios (neutron-star mergers) under consideration because in them there are conditions under which the r-process cannot be modeled without the inclusion of fission.

It is well known that heavy nuclei can undergo induced fission through neutron capture and delayed

---

*E-mail: Igor.Panov@itep.ru

fission through β-decay, and some of the nuclei with a low fission barrier can undergo spontaneous fission. Spontaneous fission seems to be of secondary importance in the r-process, because the r-process must overcome the region of nuclei $(Z > 80)$ in which the probability of induced or delayed $(Z > 88)$ fission is high before the nucleosynthesis wave reaches the region of nuclei $(Z > 92)$ for which the probability of spontaneous fission becomes significant. However, the specific contributions of different fission types to the yields of elements in the r-process can be determined only by means of consistent calculations.

On the other hand, fission can also be important for the dynamics of formation of the light fraction of heavy nuclei—through the formation from fission fragments due to the transfer of much of the matter into the actinide region by the r-process with the r-process conditions preserved.

Our goal is to calculate the rates of induced fission for the entire region of nuclei involved in the r-process in terms of a statistical model and to compare the rates of different reactions, primarily for the specific conditions of some astrophysical scenarios.

Nuclear data for fissionable and mostly deformed nuclei (fission cross sections, fission barriers, neutron binding energies, and masses of fission-product nuclei) can currently be obtained only (with a few exceptions) by means of calculations. Therefore, our second objective is to estimate the sensitivity of the results of calculations to the accuracy of the nuclear data used and attempt to estimate their reliability. It is important to emphasize that, although the mass distribution of the fission-product nuclei for long-lived fissionable nuclei has been adequately studied experimentally, the pattern of the fission-fragment mass distribution for short-lived neutron-rich actinides is unknown and studying the effect of various mass-distribution shapes on the astrophysical results (Panov *et al.* 2000) can also be useful in understanding the formation physics of fission fragments.

BINDING ENERGIES AND FISSION BARRIERS

When the nucleosynthesis wave reaches the region of actinides, new reaction channels open. These primarily include the delayed fission that was first considered by Berlovich and Novikov (1969). It competes with the emission of delayed neutrons and, depending on the nucleus characteristics, can reach 100%. If, however, the delayed-fission probability is appreciably lower than 100% and the neutron density is not high enough for the induced fission to be of great importance, then the nucleosynthesis wave can reach the region of spontaneous fission. If, alternatively, the density of free neutrons $n_n > n_{cr}$, then
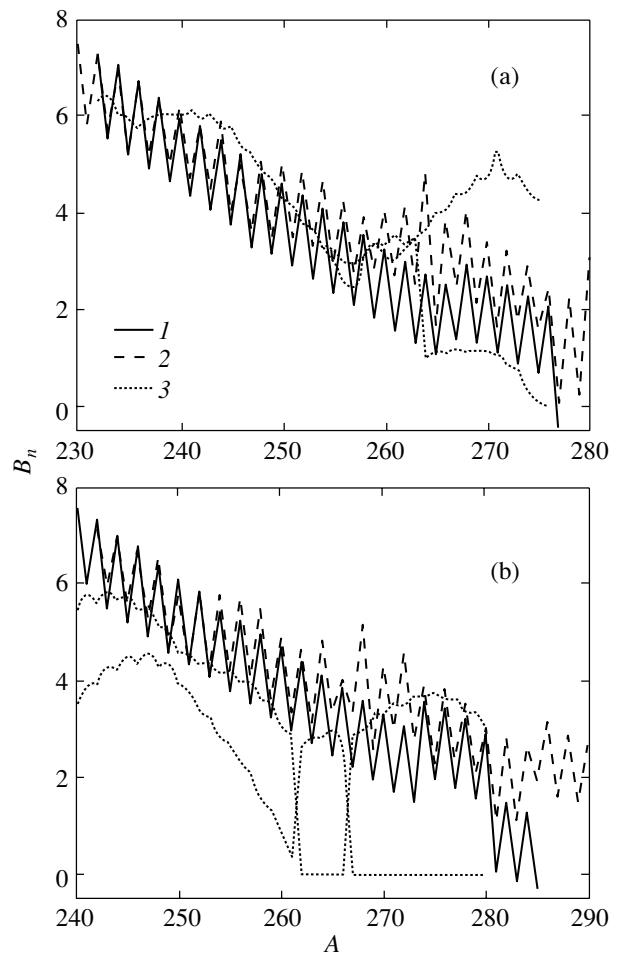
Fig. 1. The neutron binding energies calculated using different mass formulas: (*1*) from Hilf *et al.* (1976), (*2*) from Myers and Swiatecki (1998), and (*3*) the fission barriers of Howard and Möller (1980) for the isotopes of (a) uranium and (b) curium.

induced fission will dominate and spontaneous fission will be of minor importance. Although induced fission is well known and the necessity of its inclusion in the r-process was discussed previously (Kodama and Takahashi 1975; Hillebrandt and Thielemann 1977) and Thielemann *et al.* (1989) calculated the rates of induced fission for uranium isotopes, no attention was given to this process in the r-process calculations. We wish to emphasize that induced fission should be taken into account in nucleosynthesis calculations, because, as we show below, the rates of induced fission can be very high, especially for the isotopes of transuranium elements. They can exceed the β-decay rates by several orders of magnitude even near the neutron stability line. Therefore, induced fission can be a more important reaction channel than delayed fission, especially when the duration of the r-process is $\tau_r > \tau_f$.

Figure 1 shows the neutron binding energies

and fission barriers for several uranium and curium isotopes calculated using different mass formulas. One of the mass formulas (Hilf *et al.* 1976) predicts much lower neutron binding energies and, hence, smaller sizes of the r-process region, while most of the other recent mass calculations (see, e.g., Myers and Swiatecki 1998; Mamdouh *et al.* 1998) yield a weaker dependence of the neutron binding energy on the neutron excess. As a result, the neutron stability boundary passes over much more massive nuclei. This behavior is of great importance for some of the r-process scenarios, because it leads to a displacement of the r-process path into the region of shorter-lived nuclei.

In the region of actinides with mass numbers $A > 250$ under consideration, the errors in B_n and particularly in B_f can reach 1 MeV or more, which introduces a large uncertainty into the fission rates. That is why fission-barrier calculations are crucially important. Recently, new fission-barrier calculations based on the liquid-drop model (Myers and Swiatecki 1998) and on the ETFSI (Extended Thomas-Frermi + Strutinsky Integral) model have appeared (Mamdouh *et al.* 1998). The predicted barriers, especially those obtained by Mamdouh *et al.* (2001), for some isotopes are severalfold larger than the standard values of Howard and Möller (1980).

The calculated fission barriers for uranium, curium, and californium isotopes are compared in Fig. 2. As we clearly see from this figure, the dependences of the fission barrier on the neutron excess are similar in pattern for different approaches but their values differ greatly, especially in the region of isotopes with ≈ 184 neutrons. The calculations of Mamdouh *et al.* (2001) show a lack of deformation and very high fission barriers. The presented data well reflect the basic trends in the dependence of the fission barriers on the neutron excess for all actinides.

Undoubtedly, using new data in the r-process calculations can significantly change the yields of heavy elements in the region of actinides. However, it seems to us that, first, the causes of such a large discrepancy in the predicted fission barriers should be analyzed and only then should the binding energies and reactions rates reconciled with the mass formulas be used in the r-process calculations. New consistent calculations of the β-decay strength functions are required to calculate the rates of delayed fission.

FISSION CHANNELS FOR TRANSURANIUM NUCLEI

In rapid nucleosynthesis models, all three (delayed, induced, and spontaneous) fission channels were not considered simultaneously, with the probable exception of the estimates by Goriely and Clerbaux (1999). In their r-process calculations, most of the authors considered fission in a simplified form. In general, fission was assumed to take place instantly with a 100% probability when a region of extremely neutron-rich transuranium nuclei was reached. Some authors set the boundary of fissionable nuclei by choosing a limiting mass number $A > A_{\text{fis}}$ (e.g., Freiburghaus *et al.* (1999b) and Cowan *et al.* (1999) assumed that $A_{\text{fis}} = 240$ and $A_{\text{fis}} = 256$, respectively), while others (starting with Seeger *et al.* 1965) specified the boundary of the fission region through the atomic number: $Z > Z_{\text{fis}}$. These simplifications allowed the nucleosynthesis properties to be studied when the r-process was looped, but, certainly, the accuracy of calculating the yields of several nuclei, primarily cosmochronometers, decreased. In addition, because of the high delayed-fission probability, the remaining fission channels were ignored without sufficient justification, which also, at least for several scenarios, reduced the accuracy of calculating the yields of heavy nuclei.

In discussing the influence of fission on nucleosynthesis, we bear in mind that the duration of the r-process $\tau_r > \tau_f$. In this case, the nucleosynthesis wave will reach the region of transuranium elements and all fission types should be taken into account in the calculations.

In our view, spontaneous fission is not so important as induced and delayed fission for the following reasons. First, a high neutron density is required to reach the region of spontaneous fission; therefore, the rates of induced fission will be high. Second, because of the high neutron density, the delayed-fission probabilities will be close to 100% and, hence, the nucleosynthesis flux will decrease sharply toward larger Z. This was confirmed by our special calculations in which, apart from the induced and delayed fission, we took into account spontaneous fission.

The underestimation of the role of induced and delayed fission compared to spontaneous fission based on overestimated fission barriers near the neutron stability-boundary (Goriely and Clerbaux 1999) causes an appreciable increase in the cycle duration and actually reduces the allowance for fission to the simplified fission models described above.

Of course, all three fission channels should be taken into account in consistent r-process scenarios, but in this paper and our next paper (Panov and Thielemann 2003) devoted to specific calculations of the r-process in neutron-star mergers, we give particular attention to induced and delayed fission as the principal fission channels that are of greatest importance in the r-process and to their comparative role. In some of our calculations, we formally introduced spontaneous fission for nuclei with $Z > 92$ and $N > 158$. However, the inclusion of spontaneous
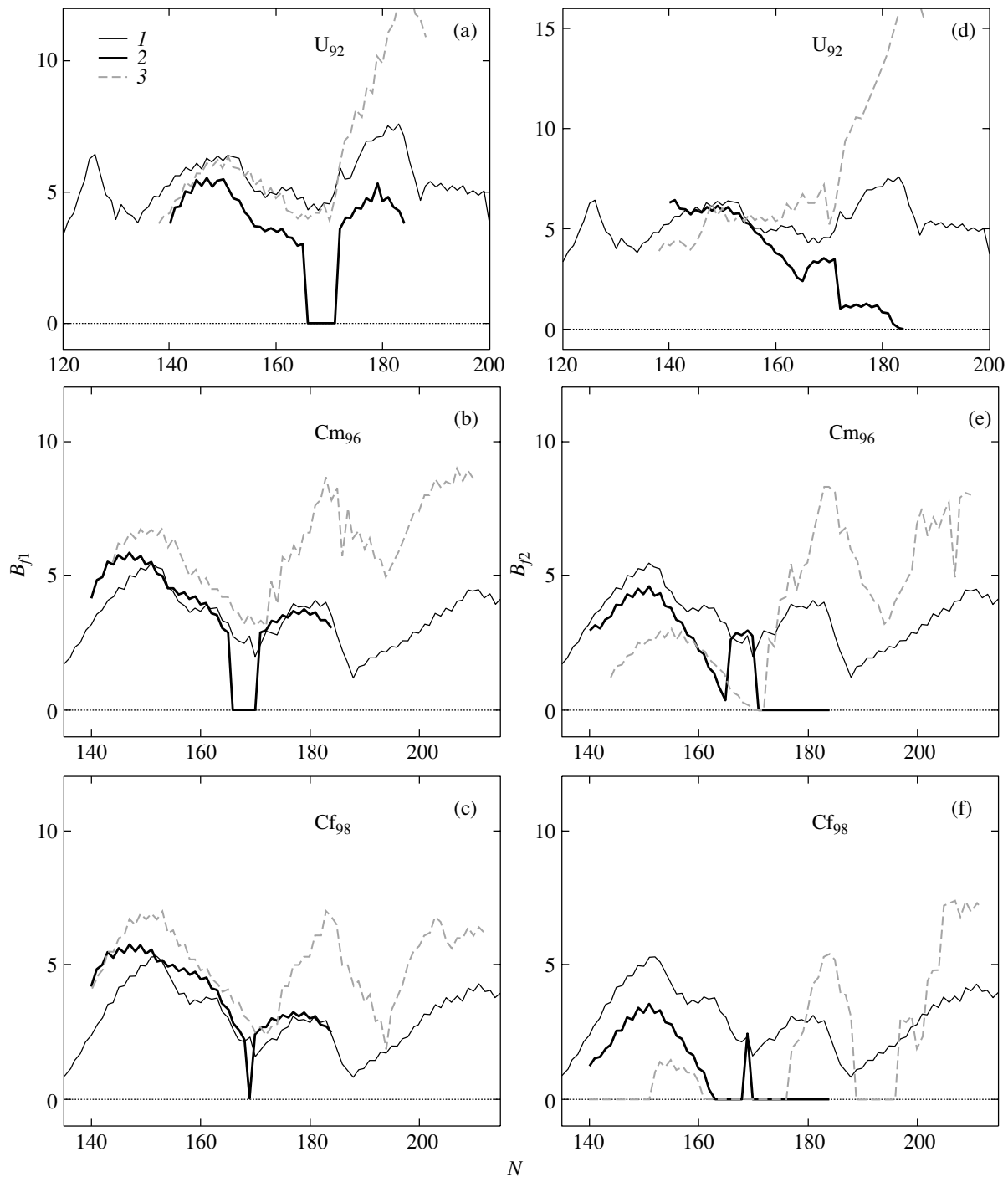
Fig. 2. The inner (a, b, c) and outer (d, e, f) fission barriers for uranium, curium, and californium isotopes calculated using different mass formulas: (*1*) from Myers and Swiatecki (1998), (*2*) from Howard and Möller (1980), and (*3*) from Mamdouh *et al.* (2001).

fission affected the formation of transuranium nuclei only slightly.

Delayed Fission

The role of delayed fission in the *r*-process was assessed long ago (see, e.g., Thielemann *et al.* 1983; Cowan *et al.* 1991; Lyutostanskiĭ *et al.* 1990). These authors determined the delayed-fission parameters from microscopic calculations and assumed that fission was most important for the formation of cosmochronometer nuclei (Cowan *et al.* 1987, 1999; Lyutostanskiĭ *et al.* 1988). It was also noticed that

this effect, at least within the scope of the classical r-process model with a nucleosynthesis duration $\tau_r \sim \tau_f$, weakly affected the yields of most heavy elements for realistic fission barriers (Cowan *et al.* 1991; Thielemann *et al.* 1989).

Such nucleosynthesis calculations in scenarios with a long r-process, $\tau_r - \tau_f \approx \tau_c$ (Rauscher *et al.* 1994; Panov *et al.* 2001), and with an allowance made for delayed fission at all rapid nucleosynthesis stages show that this fission appreciably affects the yields of the heavy elements produced in the r-process. Therefore, fission should be taken into account more accurately, particularly in the calculations of the yields of cosmochronometer nuclei.

Such characteristics of the nuclei as the reaction thresholds, delayed-process probabilities, and fission-barrier penetrability (Thielemann *et al.* 1983) must be known to calculate the delayed fission. In addition, the β-decay strength functions must be calculated. Their accuracy, as well as the accuracy of the nuclear characteristics, strongly affects the end results—the delayed-fission probabilities (Meyer *et al.* 1989; Lyutostanskii *et al.* 1990)—which gives an estimate of the maximum delayed-fission probability within the range from 30 to 100%. Although the most recent fission-barrier calculations (Mamdouh *et al.* 2001) are based on a consistent microscopic approach (Mamdouh *et al.* 1998), they give very large fission barriers for several extremely neutron-rich isotopes. Therefore, the barriers calculated by Myers and Swiatecki (1998), which also show better agreement with the experimental data (Smirenkin 1993), seem more reliable.

There are two calculations of the delayed-fission probabilities $P_{\beta df}$ for all of the nuclei for which the delayed fission is possible. They are based on two distinct theoretical models (Thielemann *et al.* 1983; Staudt and Klapdor-Kleingrothaus 1992); in these calculations, the $P_{\beta df}$ values reach 100% for a large number of nuclei but for slightly different regions of isotopes (Fig. 3).

The main difference between these calculations lies in the different models of the β-decay strength function used. The approach of Thielemann *et al.* (1983) is based on the strength-function calculations in the Tamm−Dankov approximation. In their $P_{\beta df}$ calculations, Staudt and Klapdor-Kleingrothaus (1992) used the pn-QRPA (proton−neutron quasi-particle random-phase approximation) model with an allowance made for the residual Gamow−Teller interaction to describe the strength function. In the pn-QRPA model, the states are mixed in the same way as they are in the random phase approximation (RPA). However, because of the phase factor, even a small fraction of the Gamow−Teller interaction strength

at low energies dominates over the strength of the transition to states with energies above the fission barriers, which causes the fission rates to decrease in comparison with those obtained in the earlier calculations of Thielemann *et al.* (1983).

The dependence of the delayed-fission rates on the β-decay strength function was calculated by Meyer *et al.* (1989) for the isotopes of several chemical elements. The delayed-fission probability was calculated for several tens of nuclei with various assumptions about the structure of the residual interaction and with artificially overestimated and underestimated fission barriers. It follows from the estimates that implicitly take into account the nuclear deformation (as in Staudt and Klapdor-Kleingrothaus 1992) that the calculations of Howard and Möller (1980) gave overestimated fission barriers.

For some of the protactinium isotopes ($Z = 91$), Fig. 3 shows the values of $P_{\beta df}$ calculated by Meyer *et al.* (1989) with different assumptions about the β-decay strength function (squares) and with a systematic model overestimation (triangles) and underestimation (circles) of the fission barriers. We see from this figure that the values of $P_{\beta df}$ calculated by Thielemann *et al.* (1983) were overestimated for $A < 260$ and underestimated for $A > 260$. The authors suggested that this change in the fission barriers resulted from an implicit allowance for the deformation. Their estimates showed a strong dependence of the delayed-fission probability on the fission barriers.

According to the classical model of the r-process, its path must run in the region of nuclei with a neutron binding energy of $S_n \approx 2$ MeV, for which the delayed-fission probabilities in the region of transuranium nuclei, at least for the fission barriers of Howard and Möller (1980), are highest and reach 100%. According to this view and the calculated rates of delayed fission (Thielemann *et al.* 1983), the r-process must terminate in the region of actinides with $A \approx 250-260$. Therefore, the model of instant fission with a 100% probability in the region $A \sim 260$ (for simplicity, $A = 260$) suggested by Rauscher *et al.* (1994) was a satisfactory approximation (see also similar simplifications in Cowan *et al.* (1999) and Thielemann *et al.* (1989)).

When considering the full nuclear reaction network rather than the simplified waiting-point approximation models and, in particular, in the specific scenario with rapidly changing nucleosynthesis conditions, the path of the r-process continuously changes with time (Panov 2003) and, as a result, both the yield of fission-product nuclei and the duration of the cycle change (Panov and Thielemann 2003). In particular, in the neutron-star merger model under consideration (Rosswog *et al.* 1999; Freiburghaus
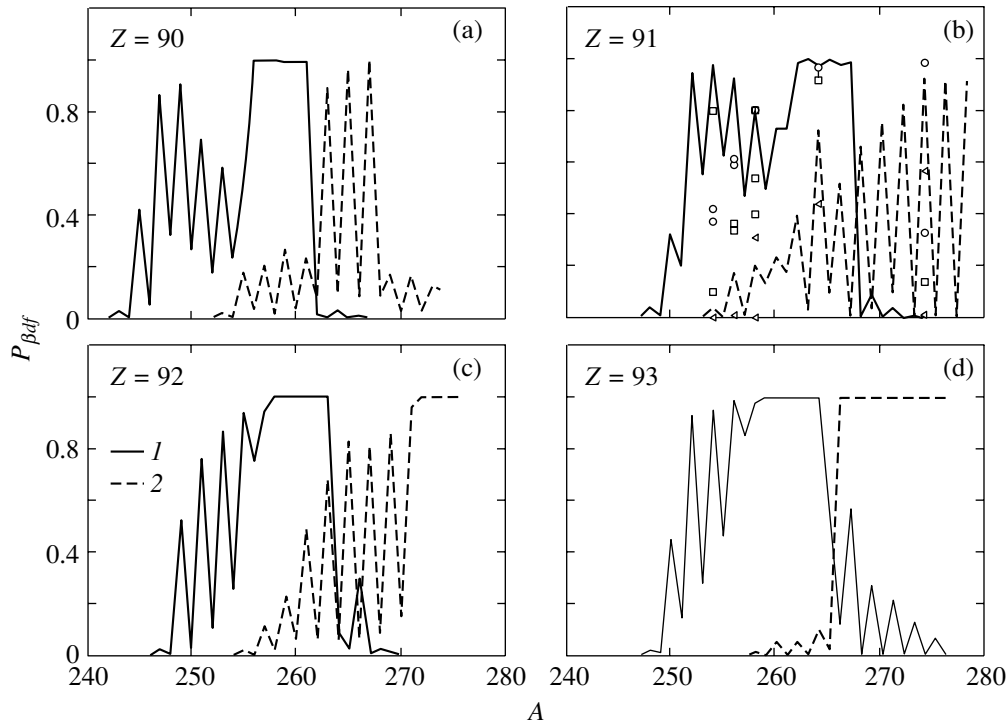
Fig. 3. The delayed-fission probability $P_{\beta df}$ versus the mass number for (a) thorium, (b) protactinium, (c) uranium, and (d) neptunium isotopes based on the calculations of Staudt and Klapdor-Kleingrothaus (1992) (*1*) and Thielemann *et al.* (1983) (*2*). Different symbols indicate the values of $P_{\beta df}$ calculated by Meyer *et al.* (1989).

et al. 1999a), the path of the r-process runs mainly along the neutron-stability boundary because of the very high density of free neutrons. In this case, for the delayed-fission rates from Thielemann *et al.* (1983), the bulk of the nucleosynthesis flux goes some distance away from the nuclei with maximum (100%) delayed-fission probabilities. Therefore, if the induced fission is disregarded, then a certain number of nuclei with $Z > 100$ and $A > 300$ can be formed.

Other calculations (Staudt and Klapdor-Kleingrothaus 1992) of the delayed-fission probabilities differ greatly from the calculations of Thielemann *et al.* (1983) mainly because a different model of the β-decay strength function was used. The main differences between the calculations can be seen from the comparison made in Fig. 3. Using different delayed-fission data leads to striking changes in the yields in both masses and chemical elements, particularly in the mass range $A < 100$. Using the fission rates of Staudt and Klapdor-Kleingrothaus (1992) can terminate the nucleosynthesis in the region of actinides under the extreme conditions achieved during neutron-star mergers when A smaller than those in the calculations with the fission rates of Thielemann *et al.* (1987) are reached.

Induced Fission

In the region of actinides, induced fission can also be an important reaction channel in the r-process. The rates of the various reactions that proceed in the r-process are compared in Fig. 4, where the rates are shown at $T = 10^9$ K. Clearly, if $\rho Y_e > 10^{-5}$, which corresponds to a neutron density $n_n > 10^{22}$, then the induced fission for nuclei with $A > 250$ will be the principal reaction channel. In this case, the nucleosynthesis wave will stop advancing into the region of even heavier nuclei because of the induced fission and the looping of the r-process, especially, as was pointed out above, when the data of Thielemann *et al.* (1983) are used. In the neutron-star merger scenario (Rosswog *et al.* 1999; Freiburghaus *et al.* 1999a), these conditions can be maintained for a long (by the r-process scales) time, $\tau \sim 0.5{-}1$ s) in highly neutronized matter.

Based on the standard statistical model of Hauser and Feshbach (1952), which is used to calculate the reaction cross sections and rates at astrophysical energies (Thielemann *et al.* 1987), we calculated the rates of induced fission for all isotopes with $Z > 80$ for the same fission barriers of Howard and Möller (1980) with which the delayed-fission rates were previously calculated by using the same model (Thielemann *et al.* 1983).
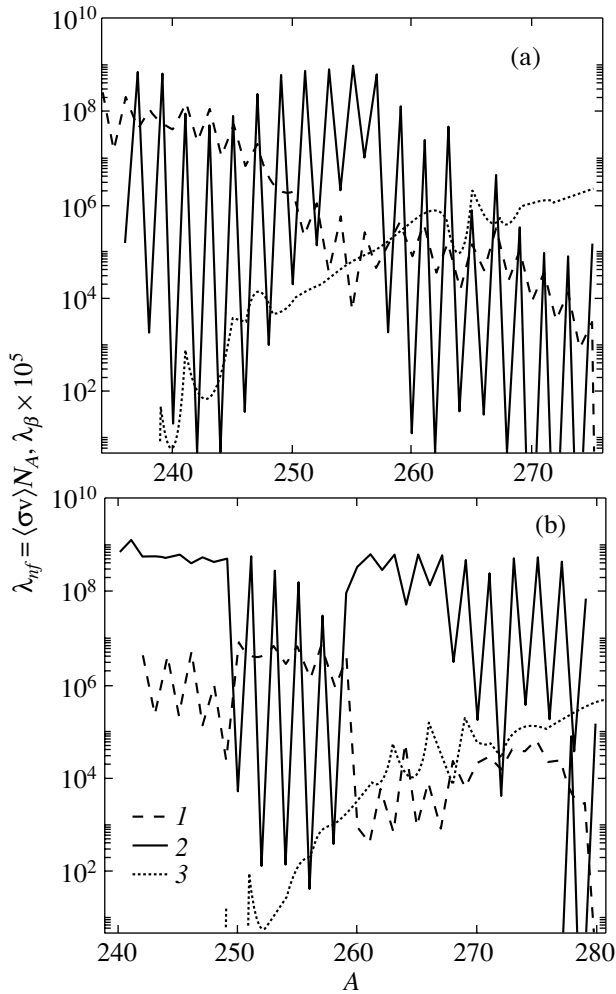
Fig. 4. The rates of neutron capture (*1*), induced fission (*2*), and β-decay (*3*) for (a) uranium and (b) californium isotopes.

The most recent calculations of Myers and Swiatecki (1998) (MS) and Mamdouh *et al.* (2001) (MPRT) indicate that the fission barriers of Howard and Möller (1980) (HM) were most likely underestimated. In some cases, the MS and MPRT fission-barrier calculations can give values that are several MeV higher than those given by the HM calculations. Therefore, in our calculations of the reaction rates, we used both the fission barriers calculated using different approaches and the HM fission barriers for uranium and curium isotopes artificially increased by 2 MeV (Tables 1–3).

It is clear from a comparison of the fission barriers calculated using the three (MS, MPRT, and HM) mass formulas given in Table 3 and shown in Fig. 2 that, in most cases, the HM fission barriers were underestimated by less than 2 MeV and only in the region of the neutron stability boundary does the difference between the HM, MS, and, particularly,

the MPRT fission barriers significantly exceed 2 MeV. This may stem from the fact that the MPRT model is more sensitive to the internal parameters than the MS model. Therefore, we also performed methodological calculations of the fission rates with the HM barriers artificially increased by 2 MeV (some of the calculations are given in Tables 1–3). A comparison of the calculations with different fission barriers shows that increasing the fission barriers for uranium isotopes by 2 MeV causes a decrease in the induced-fission rates by several orders of magnitude and a sharp decrease in the importance of induced fission for uranium isotopes in the r-process. However, as the atomic number increases, the picture slightly changes. For example, for curium isotopes, increasing the fission barriers also causes the fission rates to decrease, but they continue to be large and the fission of neutron-rich curium isotopes in the r-process will remain a major reaction channel. The influence of errors in the fission-barrier calculations on the element yields in the r-process will be discussed in our next paper (Panov and Thielemann 2003).

We believe that induced fission is particularly important in the cycling of the r-process when the density of free neutrons is still sufficiently high. In this case, the rates of induced fission are so high (Tables 1–3; see also Thielemann *et al.* 1989) that this reaction channel will continue to be important even with the higher fission barriers obtained recently (Mamdouh *et al.* 2001).

The Mass Distribution of Fission-Product Nuclei

The mass distribution of fission-product nuclei is a separate question. There are at least two important consequences of the pattern of mass distribution. First, if the total yield of the fission fragments is strongly asymmetric, then a significant number of isotopes with masses $90 < A < 120$ can be produced. Second, if $\tau_r \approx \tau_f$, then different representations of the mass distribution and, hence, different neutron yields can also affect the yields of the isotopes for masses $130 < A < 190$ and actinides at the final stage of the r-process (Panov and Thielemann 2003).

Cowan *et al.* (1999) and Rauscher *et al.* (1994) assumed that the mass distribution during delayed fission could be only symmetric. If we consider the classical model of the r-process with the nucleosynthesis proceeding along nuclei with a binding energy of ≈ 2 MeV and if the delayed-fission probabilities are close to those calculated by Thielemann *et al.* (1983), then the symmetric distribution is a good approximation, because the path of the r-process runs precisely along the nuclei with $A \approx 260$. Besides, the delayed-fission probability for these nuclei is at a maximum.

Table 1. The reaction rates $\langle \sigma v \rangle N_A$ for the neutron-rich uranium isotopes produced in the r-process

$A(Z=92)$	S_n	B_{f1}	B_{f2}	$\lambda_{n,\gamma}$	$\lambda_{n,f}(B_{fi})$	$\lambda_{n,f}(B_{fi}+2)$	$P_{\beta df}$	λ_β
248	3.22	3.69	4.86	0.168E+07	9.96D+02	3.21D−08	0.020	0.569E−01
249	4.60	3.67	4.72	0.948E+06	6.25D+08	7.30D+01	0.010	0.468E−01
250	2.95	3.52	4.30	0.943E+06	1.99D+04	8.28D−07	0.440	0.947E−01
251	4.32	3.64	4.18	0.992E+05	7.84D+08	8.06D+02	0.140	0.138E+00
252	2.67	3.47	3.79	0.534E+06	1.43D+05	1.37D−05	0.930	0.179E+00
253	4.04	3.60	3.69	0.187E+05	8.39D+08	8.36D+03	0.280	0.310E+00
254	2.40	3.33	3.27	0.302E+06	2.11D+06	1.42D−03	0.950	0.423E+00
255	3.77	3.28	3.03	0.308E+04	8.42D+08	1.76D+05	0.460	0.579E+00
256	2.13	2.97	2.58	0.123E+06	1.23D+07	4.93D−02	0.980	0.778E+00
257	3.49	3.00	2.40	0.209E+05	6.42D+08	2.46D+02	0.850	0.128E+01
258	1.86	0.0	3.09	0.855E+05	1.90D+03	6.18D−08	0.980	0.210E+01
259	3.22	0.0	3.37	0.226E+06	1.42D+08	1.28D+00	1.000	0.164E+01
260	1.60	0.0	3.32	0.414E+05	1.32D+01	4.22D−10	1.000	0.475E+01
261	2.96	0.0	3.54	0.189E+06	2.70D+07	5.23D−02	1.000	0.597E+01
262	1.34	0.0	3.36	0.184E+05	1.11D+00	3.38D−11	1.000	0.820E+01
263	2.70	0.0	3.50	0.789E+05	5.27D+07	5.71D+02	1.000	0.461E+01
264	1.11	3.61	1.01	0.806E+04	4.08D+01	5.14D−06	1.000	0.861E+00
265	2.52	4.03	1.17	0.716E+05	8.43D+05	5.04D+00	0.480	0.251E+02
266	1.40	4.02	1.10	0.193E+05	3.35D+01	7.25D−06	0.130	0.658E+01
267	2.90	4.41	1.21	0.159E+06	4.75D+06	4.08D+01	0.570	0.694E+01
268	1.33	4.38	1.12	0.122E+05	1.17D+00	2.69D−07	0.010	0.382E+01
269	2.69	4.75	1.26	0.583E+05	3.83D+05	2.50D+00	0.270	0.101E+02
270	1.12	4.65	1.11	0.446E+04	1.81D−03	4.96D−10	0.010	0.105E+02
271	2.48	5.33	1.18	0.206E+05	1.05D+05	7.21D−01	0.210	0.142E+02
272	0.92	4.66	0.83	0.163E+04	9.05D−03	5.46D−09	0.010	0.115E+02
273	2.27	4.79	0.79	0.798E+04	9.46D+04	6.77D−01	0.130	0.149E+02
274	0.72	4.44	0.29	0.408E+03	5.92D−02	1.61D−08	0.000	0.184E+02
275	2.07	4.28	0.08	0.000E+00	1.69D+05	1.72D+00	0.070	0.228E+02

However, for more realistic nucleosynthesis scenarios, as in the case where the delayed-fission probabilities differ from those calculated by Thielemann *et al.* (1983), most of the masses of the fissionable nuclei are not equal to 260. The mass distribution of the fission-product nuclei can then be asymmetric, which may lead to significant yields of the isotopes in the mass range $80 < A < 120$.

In the most general case, the ratio of the fission-product yields through different (symmetric or asymmetric) fission modes of transuranium nuclei can be written as

$$R_{\beta f}^{\text{fis}} = \frac{\sum_{A_i, Z_j} P_{\beta f}^{\text{asym}} Y_{A_j, Z_j} \lambda_\beta^{ij}}{\sum_{A_i, Z_j} P_{\beta f}^{\text{symm}} Y_{A_j, Z_j} \lambda_\beta^{ij}}. \tag{1}$$

This ratio $R_{\beta f}^{\text{fis}}$ strongly depends primarily on the branching ratios of the most abundant nuclei that are produced during the nucleosynthesis and, hence, on the nuclear data used.

Since the mass distribution for the short-lived nu-

Table 2. The reaction rates $\langle \sigma v \rangle N_A$ for the neutron-rich curium isotopes produced in the r-process

$A(Z=96)$	S_n	B_{f1}	B_{f2}	$\lambda_{n,\gamma}$	$\lambda_{n,f}(B_{fi})$	$\lambda_{n,f}(B_{fi}+2)$	$P_{\beta df}$	λ_β
251	5.74	4.53	3.80	0.922E+07	6.58D+08	6.58D+08	0.000	0.504E−04
252	4.09	4.29	3.36	0.106E+08	1.47D+02	1.47D+02	0.000	0.292E−04
253	5.46	4.37	3.25	0.180E+08	3.19D+08	3.19D+08	0.000	0.144E−03
254	3.81	4.11	2.80	0.624E+07	1.59D+02	1.59D+02	0.000	0.321E−03
255	5.18	4.22	2.69	0.190E+08	1.78D+08	1.78D+08	0.000	0.148E−02
256	3.54	3.92	2.24	0.366E+07	4.82D+01	4.82D+01	0.010	0.250E−02
257	4.90	3.98	2.08	0.213E+08	3.30D+07	3.30D+07	0.000	0.925E−02
258	3.27	3.59	1.60	0.216E+07	4.25D+02	4.25D+02	0.080	0.119E−01
259	4.63	3.53	1.34	0.938E+07	1.02D+08	1.02D+08	0.020	0.206E−01
260	3.00	3.04	0.79	0.246E+04	3.87D+08	6.51D+04	0.810	0.330E−01
261	4.35	2.88	0.40	0.101E+04	7.12D+08	5.85D+07	0.330	0.904E−01
262	2.74	0.0	2.62	0.193E+05	3.23D+08	7.15D−01	1.000	0.877E−01
263	4.08	0.0	2.86	0.205E+04	7.01D+08	2.55D+06	0.990	0.691E+00
264	2.47	0.0	2.78	0.130E+06	6.00D+07	8.51D−03	1.000	0.121E+00
265	3.82	0.0	2.97	0.229E+04	6.91D+08	2.17D+05	1.000	0.142E+00
266	2.21	0.0	2.76	0.253E+05	1.49D+08	2.61D+02	0.990	0.199E+01
267	3.55	2.88	0.0	0.210E+04	6.83D+08	5.58D+06	0.990	0.649E+00
268	1.96	2.96	0.0	0.619E+05	4.07D+06	1.35D+00	0.970	0.844E−01
269	3.29	3.31	0.0	0.178E+05	5.84D+08	2.15D+05	0.990	0.284E+01
270	1.71	3.26	0.0	0.587E+05	2.00D+05	4.22D−02	0.790	0.602E+00
271	3.05	3.52	0.0	0.794E+05	2.99D+08	1.48D+04	0.970	0.714E+00
272	1.54	3.44	0.0	0.429E+05	5.03D+03	8.52D−04	0.450	0.332E+00
273	3.48	3.70	0.0	0.120E+06	6.48D+08	1.84D+05	0.900	0.126E+01
274	2.02	3.57	0.0	0.108E+06	4.23D+05	9.60D−02	0.420	0.153E+01
275	3.36	3.76	0.0	0.190E+06	6.59D+08	8.24D+04	0.970	0.153E+01
276	1.81	3.54	0.0	0.545E+05	2.09D+05	5.67D−02	0.870	0.130E+01
277	3.15	3.64	0.0	0.714E+05	5.49D+08	7.81D+04	0.990	0.205E+01
278	1.61	3.35	0.0	0.953E+04	4.30D+04	1.78D−02	0.710	0.290E+01

clei under consideration is not known, we considered several different models.

In our calculations, we assumed that the fission is completely symmetric only for the region of nuclei with masses $A > 255$. For the nuclei with $A \approx 260$, this symmetry agrees with the observations, while the mass distribution of the high-mass fissionable nuclei is unknown. In other cases ($A < 255$ and for $Z < 90$), the fission was assumed to be asymmetric.

Based on the available experimental and theoretical data (Itkis *et al.* 1989; Schmidt *et al.* 2001),

we considered two simple models of the mass distribution for the fission products (Panov and Thielemann 2003).

In the case of asymmetric fission, the masses and charges of the fission products (the subscripts "H" and "L" denote heavy and light fragments, respectively) were defined as follows (model M1):

$$A_H = 0.57\,A, \quad A_L = 0.43\,A, \qquad (2)$$
$$Z_H = 0.57\,Z, \quad Z_L = 0.43\,Z.$$

Table 3. The induced-fission rates $\langle \sigma v \rangle N_A$ for some of the uranium and curium isotopes calculated with various fission-barrier models

Element	Models	250	255	260	265	270	275	280	285
U	HM	1.99D4	8.42D8	1.32d1	8.43d5	1.81d−3	1.69d5	−	−
U	HM+2 MeV	8.28d−7	1.76d5	4.22d−10	5.04	4.96d−10	1.72	−	−
U	MS	1.3d3	2.14d7	13.9	1.3d5	3.65d−7	7.69d−3	6.97d−4	948.
U	MPRT	8.0d−5	396	4.8d−14	4.d−18	7.7d−43	0	−	−
Cm	HM	2.86d3	1.78d8	3.87d8	6.91d8	2.0d5	6.59d8	−	−
Cm	HM+2 MeV	2.86d3	1.78d8	6.51d4	2.17d5	4.22d−2	8.24d4	−	−
Cm	MS	5.09d8	7.63d8	3.88d8	6.93d8	3.14d8	9.63d8	2.37d4	9.3d8
Cm	MPRT	1.35d4	7.64d8	2.17d6	6.93d8	31.4	332	2.7d−27	3.2d−3

In the neutron-star merger scenario, we also considered two fission modes, as in the schematic model of the r-process, but the charge and mass numbers of the fragments were defined as prescribed by Itkis et al. (1989) (model M2):

$$A_H = 130, \quad A_L = A - A_H, \qquad (3)$$
$$Z_H = 52 - (Z - 80)/10, \quad Z_L = A - A_H.$$

Most of our calculations were performed with different delayed-fission rates (Thielemann et al. 1983; Staudt and Klapdor-Kleingrothaus 1992) for the mass distribution of the fission fragments specified by model M2, which gave us preliminary information on the influence of the fission rates (and the contribution of different fission modes) on the final elemental abundances. The ratio $R_{\beta f}^{\text{fis}}$ integrated over time and over all nuclei for the model of Thielemann et al. (1983) in our calculations with different assumptions about the mass distribution of the fission products suggests approximately equal contributions of the symmetric and asymmetric modes to the abundances of the final (mostly secondary) nuclei. For this model, $R_{\beta f}^{\text{fis}}$ for the region of nuclei with $Z \leq 96$ (from which the nucleosynthesis does not emerge) is 0.5 (in the neutron-star merger scenario) and 0.75 (in the simple schematic model of the r-process) when the delayed-fission reaction rates from Thielemann et al. (1983) are used. For other delayed-fission data (Staudt and Klapdor-Kleingrothaus 1992), the relative contribution of the symmetric and asymmetric fission will be different, more specifically, $R_{\beta f}^{\text{fis}} \approx 1.9$.

A comparison of the results of our calculations indicates that the shape of the mass distribution (M1 and M2) for the data of Staudt and Klapdor-Kleingrothaus (1992) weakly affects the parameter $R_{\beta f}^{\text{fis}}$ for the nuclei with $N > 155$ and $Z > 92$.

Although the contributions of the symmetric and asymmetric fission are comparable, specific numbers can vary, because we only took into account delayed fission in our calculations of $R_{\beta f}^{\text{fis}}$.

We analyzed in detail the influence of the mass distribution of the fission-product nuclei on the yields of various nuclei in nucleosynthesis, in particular, the cosmochronometer nuclei, the knowledge of whose abundances is of great importance in determining the age of the Galaxy, in a separate paper (Panov and Thielemann 2003).

CONCLUSIONS

The delayed- and induced-fission rates are compared in Fig. 5, where the circles and squares of different sizes represent different fission rates. Clearly, the delayed fission for $Z \leq 87$ is the principal fission mode. In the range $Z = 88-90$, induced and delayed fission compete with each other, while for the isotopes of the chemical elements with $Z > 91$, the induced fission dominates. (Of course, the production rate of secondary nuclei through induced fission also strongly depends on the density of free neutrons at the r-process site.) The open squares represent the nuclei with a low delayed-fission rate. This is possible for the region of nuclei with a large neutron excess ($N > 160$), because the neutron binding energies are low ($S_n < B_f$) and the daughter nucleus decays mainly through neutron evaporation rather than through fission. The path of the r-process (Panov 2003) that is determined by the most abundant isotopes of each chemical element ($(A, Z) > 1\%Y(Z)$) is indicated by crosses and pluses at the beginning (when $n_n > 10^{26}$ cm^{-3}) and end (when the neutron density is less than 10^{19} cm^{-3}) of the r-process, respectively.
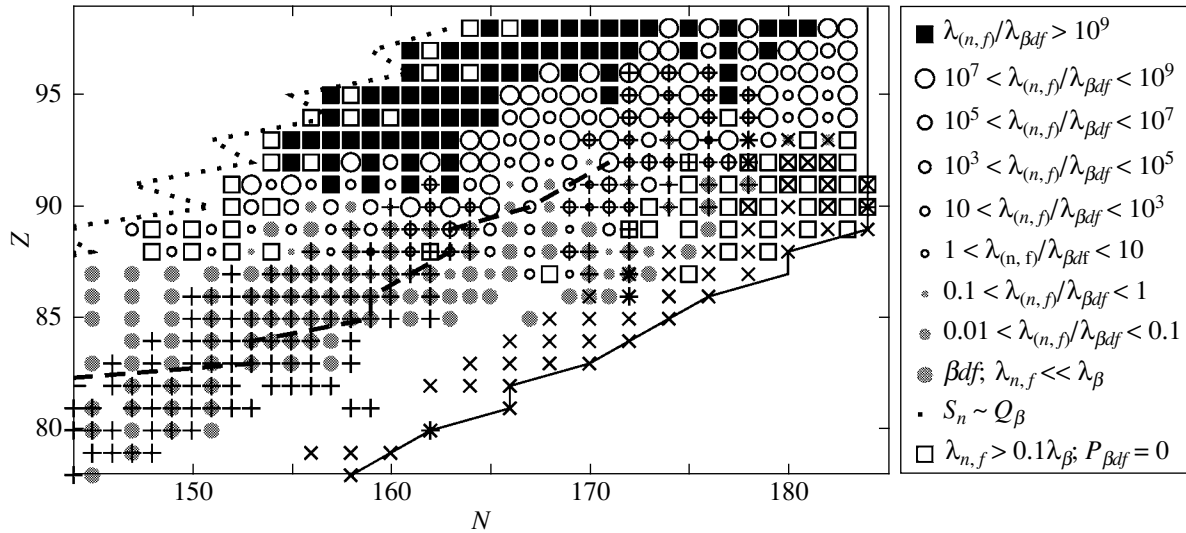
Fig. 5. A map of the rate ratio $\lambda_{n,f}/\lambda_{\beta df}$. The most abundant nuclei along the path of the r-process are represented by crosses (when $n_n \geq 10^{26}$) or pluses (when n_n falls below $\sim 10^{19}$); the lines indicate the neutron-stability boundary (solid line), the positions of the nuclei with a neutron binding energy $S_n \approx 2$ MeV (dashed line), and the boundary of the nuclei with β-decay energies $Q_\beta \sim S_n$ (dotted line).

It follows from our analysis of the influence of the fission barriers on the reaction rates that not only delayed fission but also induced fission can be of great importance under the conditions of naturally occurring rapid nucleosynthesis scenarios. The induced-fission rates at astrophysical energies have been calculated for the first time for a large number of isotopes (about 400). Thus, we can compute the r-process by taking into account not only delayed fission but also induced fission and determine their comparative role during nucleosynthesis. Our estimates also indicate that an increase in the fission barriers, as suggested by the most recent calculations (Myers and Swiatecki 1998; Mamdouh *et al.* 2001), leads not so much to a decrease in the importance of fission as to the possible synthesis of heavier elements than those for the Howard–Möller fission barriers (Panov and Thielemann 2003) in the r-process.

ACKNOWLEDGMENTS

REFERENCES

1. E. Yu. Berlovich and Yu. N. Novikov, Dokl. Akad. Nauk SSSR **185**, 1025 (1969) [Sov. Phys. Dokl. **14**, 349 (1969)].
2. J. J. Cowan, F.-K. Thielemann, and J. W. Truran, Astrophys. J. **323**, 543 (1987).
3. J. J. Cowan, F.-K. Thielemann, and J. W. Truran, Phys. Rep. **208**, 267 (1991).
4. J. J. Cowan, B. Pfeiffer, K.-L. Kratz, *et al.*, Astrophys. J. **521**, 194 (1999).
5. C. Freiburghaus, S. Rosswog, and F.-K. Thielemann, Astrophys. J. Lett. **525**, L121 (1999a).
6. C. Freiburghaus, J.-F. Rembges, T. Rauscher, *et al.*, Astrophys. J. **516**, 381 (1999b).
7. S. Goriely and B. Clerbaux, Astron. Astrophys. **346**, 798 (1999).
8. W. Hauser and H. Feshbach, Phys. Rev. **78**, 357 (1952).
9. E. R. Hilf, H. V. Groote, and K. Takahashi, CERN-report. 1976. CERN-76-13, p. 142.
10. V. Hill, B. Plez, R. Cayrel, *et al.*, Astron. Astrophys. **387**, 560 (2002).
11. W. Hillebrandt and F.-K. Thielemann, Astron. Astrophys. **58**, 357 (1977).
12. W. M. Howard and P. Möller, At. Data Nucl. Data Tables **25**, 219 (1980).
13. M. G. Itkis, V. N. Okolovich, and G. N. Smirenkin, Nucl. Phys. A **502**, 243 (1989).
14. T. Kodama and K. Takahashi, Nucl. Phys. A **239**, 489 (1975).
15. Yu. S. Lyutostanskii, S. V. Malevannyi, and I. V. Panov, Yad. Fiz. **47**, 780 (1988).
16. Yu. S. Lyutostanskii, I. V. Panov, and V. I. Lyashuk, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 2137 (1990).

17. A. Mamdouh, J. M. Pearson, M. Rayet, and F. Tondeur, Nucl. Phys. A **644**, 389 (1998).

18. A. Mamdouh, J. M. Pearson, M. Rayet, and F. Tondeur, Nucl. Phys. A **679**, 337 (2001).

19. B. S. Meyer, W. M. Howard, G. J. Mathews, *et al.*, Phys. Rev. C **39**, 1876 (1989).

20. W. D. Myers and W. J. Swiatecki, Phys. Rev. T **60**, 014606 (1998).

21. I. V. Panov, Pis'ma Astron. Zh. **29**, 195 (2003) [Astron. Lett. **29**, 163 (2003)].

22. I. V. Panov and V. M. Chechetkin, Pis'ma Astron. Zh. **28**, 541 (2002) [Astron. Lett. **28**, 476 (2002)].

23. I. V. Panov and F.-K. Thielemann, *Nuclei in the Cosmos VII,* Ed. by S. Kubono (Fuji-Yoshida, Japan, 2003), in press.

24. I. V. Panov, C. Freiburghaus, and F.-K. Thielemann, *Workshop on Nuclear Astrophysics,* Ed. by W. Hillebrandt and E. Müller (MPA, Ringberg, Garching, 2000), p. 73.

25. I. V. Panov, T. Freiburghaus, and F.-K. Thielemann, Nucl. Phys. A **688**, 587 (2001).

26. Y.-Z. Qian and G. J. Wasserburg, Astrophys. J. **567**, 515 (2002).

27. T. Rauscher, J. H. Applegate, J. J. Cowan, *et al.*, Astrophys. J. **429**, 499 (1994).

28. S. Rosswog, M. Liebendörfer, F.-K. Thielemann, *et al.*, Astron. Astrophys. **341**, 499 (1999).

29. K.-H. Schmidt, J. Benlliure, C. Böckstiegel, *et al.*, Nucl. Phys. A **685**, 60 (2001).

30. P. A. Seeger, W. A. Fowler, and D. D. Clayton, Astrophys. J., Suppl. **11**, 121 (1965).

31. G. N. Smirenkin, IAEA Report INDC(CCP)-359 (1993).

32. C. Sneden, J. J. Cowan, I. I. Ivans, *et al.*, Astrophys. J. Lett. **533**, L139 (2000).

33. A. Staudt and H. V. Klapdor-Kleingrothaus, Nucl. Phys. A **549**, 254 (1992).

34. F.-K. Thielemann, M. Arnould, and W. Truran, *Advances in Nucl. Astrophysics,* Ed. by E. Vangioni-Flam *et al.* (Frontiers, Gif sur Yvette, 1987), p. 525.

35. F.-K. Thielemann, A. G. W. Cameron, and J. J. Cowan, *Int. Conf. 50 Years with Nuclear Fission,* Ed. by J. Behrens and A. D. Carlson (Americ. Nucl. Soc., La Grange Park, 1989), p. 592.

36. F.-K. Thielemann, J. Metzinger, and H. V. Klapdor-Kleingrothaus, Z. Phys. A **309**, 301 (1983).

37. G. J. Wasserburg, M. Busso, and R. Gallino, Astrophys. J. Lett. **466**, L109 (1996).

Translated by V. Astakhov

Radial Pulsations of Helium Stars with Masses from 1 to 10 $M_\odot$

Yu. A. Fadeyev[*] and M. F. Novikova

Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia
Received March 5, 2003

Abstract—We present the results of our hydrodynamic calculations of radial pulsations in helium stars with masses $1\,M_\odot \leq M \leq 10\,M_\odot$, luminosity-to-mass ratios $1 \times 10^3 L_\odot/M_\odot \leq L/M \leq 2 \times 10^4 L_\odot/M_\odot$, and effective temperatures 2×10^4 K $\leq T_{\rm eff} \leq 10^5$ K for mass fractions of helium $Y = 0.98$ and heavy elements $Z = 0.02$. We show that the lower boundary of the pulsation-instability region corresponds to $L/M \sim 10^3 L_\odot/M_\odot$ and that the instability region for $L/M \lesssim 5 \times 10^3 L_\odot/M_\odot$ is bounded by effective temperatures $T_{\rm eff} \lesssim 3 \times 10^4$ K. As the luminosity rises, the instability boundary moves into the left part of the Hertzsprung−Russell diagram and radial pulsations can arise in stars with effective temperatures $T_{\rm eff} \lesssim 10^5$ K at $L/M \gtrsim 7 \times 10^3 L_\odot/M_\odot$. The velocity amplitude for the outer boundary of the hydrodynamic model increases with L/M and lies within the range $200 \lesssim \Delta U \lesssim 700$ km s^{-1} for the models under consideration. The periodic shock waves that accompany radial pulsations cause a significant change of the gas-density distribution in the stellar atmosphere, which is described by a dynamic scale height comparable to the stellar radius. The dynamic instability boundary that corresponds to the separation of the outer stellar atmospheric layers at a superparabolic velocity is roughly determined by a luminosity-to-mass ratio $L/M \sim 3 \times 10^4 L_\odot/M_\odot$. © 2003 MAIK "Nauka/Interperiodica".

Key words: *stars—variable and peculiar, stellar pulsations, helium stars, and Wolf−Rayet stars.*

INTRODUCTION

The ionization of iron-group elements at a gas temperature $T \sim 2 \times 10^5$ K is accompanied by a significant increase in the opacity of matter and can lead to a pulsation instability of the star. In particular, early-type helium stars are pulsating variables with this instability-excitation mechanism. Until recently, radial pulsations have been observed in the variables V652 Her and BX Cir, which belong to the group of low-mass ($M \lesssim 1\,M_\odot$, $L \sim 10^3 L_\odot$) helium stars. However, since pulsations have been detected in some of the Wolf−Rayet stars (Monderen *et al.* 1988; Bratschi and Blecha 1996; Marchenko *et al.* 1998; Morel *et al.* 1999; Veen *et al.* 2002a, 2000b, 2000c), analysis of the pulsation instability of high-mass helium stars arouses particular interest.

A linear analysis of the pulsation instability indicates that both radial and nonradial pulsation modes can be excited in high-mass early-type helium stars (Glatzel and Mehren 1996; Glatzel *et al.* 1993; Gautschy 1995; Kiriakidis *et al.* 1996). This circumstance seems to be of crucial importance, because the periodic shock waves produced by pulsations cause the gas density in the outer stellar atmospheric layers to significantly increase. In turn, the assumption of

a higher gas density near the sonic point makes it possible to bring the calculations of a steady-state stellar wind into much better agreement with observational data (Heger and Langer 1996) and to explain the fact that the momentum of the outflowing gas in Wolf−Rayet stars always exceeds the photon momentum (Heger and Langer 1996; Hamann and Koesterke 1998; Herald *et al.* 2000; Nugis and Lamers 2000): $\dot{M} v_\infty c/L > 1$, where $\dot{M}$ is the mass-loss rate and v_∞ is the terminal velocity of the gas flow.

Among several studies (Glatzel *et al.* 1999; Cox and Cahn 1988; Maeder 1985; Schaller 1990) in which the pulsation instability of Wolf−Rayet stars was modeled, only the calculations of Glatzel *et al.* were carried out using new data on the opacity and the equation of state of stellar matter. The three models of helium stars with masses of 6, 14, and 64 $M_\odot$ considered by these authors showed a radial-mode instability with a gas-velocity amplitude at the outer boundary of the model $\Delta U \gtrsim 10^2$ km s^{-1}, indicating the necessity of further work in this direction.

In this paper, we analyze the radial pulsations of helium stars. Compared to the nonradial oscillations, the spherically symmetric gas motions are only a special case of stellar pulsations. Nevertheless, the necessity of analyzing the spherically symmetric motions stems from the fact that the radial pulsations
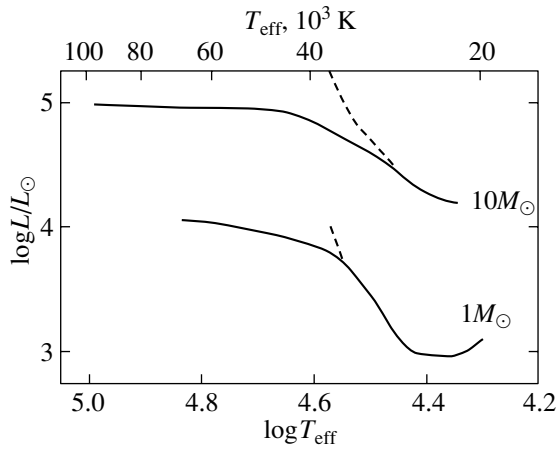
Fig. 1. The instability boundaries of radially pulsating helium stars with masses of 1 and 10 $M_\odot$. The dashed lines represent the boundaries of the fundamental-mode pulsation region.

Fig. 2. (a) The normalized radial displacement amplitude $\Delta R/R$ at the outer boundary of the model versus the effective temperature $T_{\rm eff}$; (b) a similar dependence for the velocity amplitude ΔU. The luminosities L are given near the curves. The individual models are represented by the filled circles.

have a much larger amplitude than the nonradial oscillations and, hence, they prove to be an incomparably more efficient mass-loss mechanism. Moreover, the radial pulsations can be analyzed much more thoroughly by using the nonlinear theory. In particular, nonlinear calculations allow us to determine the amplitude of the motions and estimate the extent to which the stellar pulsations affect the gas density distribution in the stellar atmosphere.

THE MODEL

The results of calculations presented below cover the mass range $1 \leq M/M_\odot \leq 10$. Thus, they supplement our previously computed (Fadeyev and Novikova 2003) grid of model helium stars with masses $0.5\,M_\odot \leq M \leq 0.9\,M_\odot$. The only significant distinction compared to our previous work is that we used the OPAL tables (Rogers *et al.* 1996; Iglesias and Rogers 1996) to calculate the equation of state and opacity of the stellar gas. We used these data instead of the OP tables (Seaton *et al.* 1994) that we used previously because the OPAL tables cover wider temperature and density ranges. This circumstance is of particular importance in considering helium stars with effective temperatures $T_{\rm eff} \gtrsim 3 \times 10^4$ K. Moreover, according to the calculations of Iglesias and Rogers (1995), the OP opacities underestimate the atomic level populations at high temperatures, which results in smaller photoionization cross sections and, hence, in underestimated mean Rosseland opacities.

As in our previous paper (Fadeyev and Novikova 2003), here we consider the self-excited, spherically symmetric stellar pulsations that develop from a hydrostatic and thermal equilibrium under the
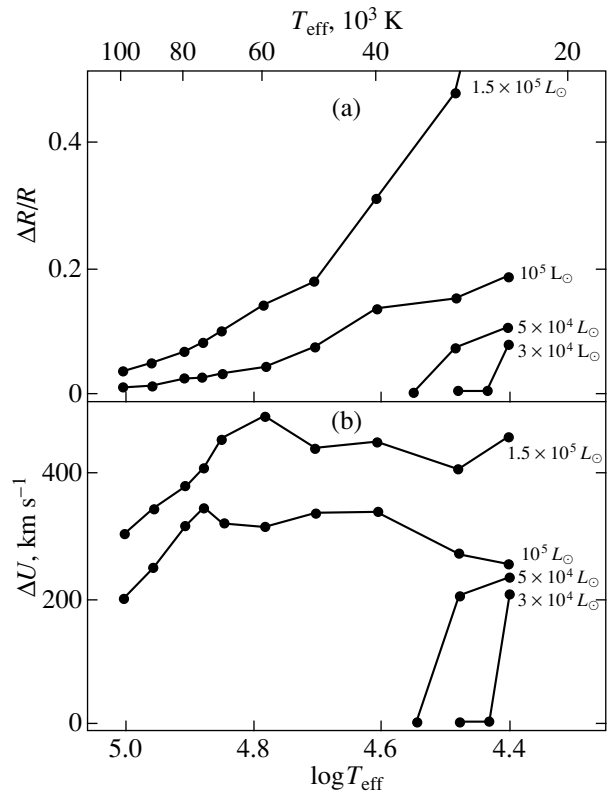
effect of small initial perturbations. The computational noise, i.e., the approximation errors related to the finite-difference representation of the differential equations, acts as such perturbations. The smallness of these errors follows from the fact that the initial kinetic energy of the stellar envelope is several orders of magnitude lower than the limiting value that corresponds to the regime of self-excited oscillations. Moreover, in limit-cycle models, the values of the hydrodynamic variables in all of the Lagrangian zones are repeated in every oscillation cycle with a relative error $\varepsilon \lesssim 10^{-4}$.

Since we used from 200 to 500 Lagrangian intervals for the finite-difference representation of our hydrodynamic models, from 10^4 to 5×10^4 integration steps in time t were required to compute one pulsation cycle. Most of the models (more than 200) were computed for helium and heavy-element mass fractions of $Y = 0.98$ and $Z = 0.02$, respectively. To assess the role of heavy elements in the growth of pulsation instability, we computed several tens of models for $Y = 0.99$, $Z = 0.01$ and $Y = 1$, $Z = 0$. We analyzed the hydrodynamic models that showed strictly repet-
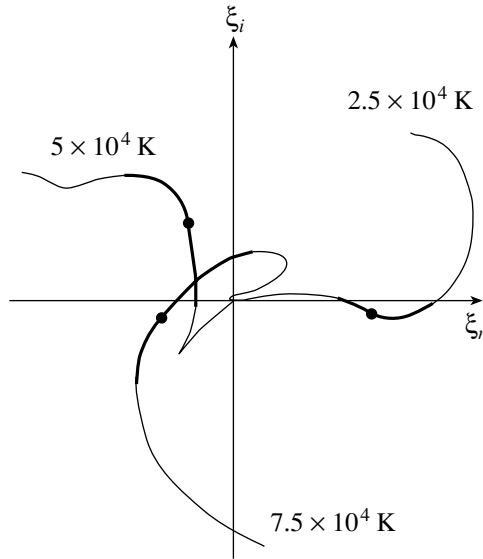
Fig. 3. Normalized phase diagrams of the fundamental Fourier harmonic for model helium stars with $M = 10\,M_\odot$, $L = 10^5\,L_\odot$, and $Z = 0.02$. The bold segments of the curves represent the excitation regions of the pulsation instability, and the filled circles indicate the layer with maximum mechanical work per cycle. The effective temperatures are given near the curves.

itive limit–cycle oscillations by the Fourier decomposition of the solution (Fadeyev 1994) and used the discrete Fourier transform of the velocity U and radius r of all Lagrangian layers in the time interval t spanning from 10^2 to 10^3 pulsation cycles for the models with irregular oscillations (Fadeyev 1998, 2000).

THE PULSATION-INSTABILITY STRIP IN THE HERTZSPRUNG–RUSSELL DIAGRAM

We can get a general idea of the properties of pulsating helium stars by examining the pulsation-instability region in the Hertzsprung–Russell (HR) diagram. At fixed values of M, Y, and Z, the boundary of this region is the line of zero instability growth rate $\eta = \Pi d\ln E_{\mathrm{K}}/dt = 0$, where Π is the pulsation period and E_{K} is the kinetic energy of the pulsating envelope. The $\eta = 0$ lines for helium stars with masses of 1 and $10\,M_\odot$ are shown in Fig. 1. Since the growth rate η increases with stellar luminosity L, the stars that are unstable against radial oscillations are located in the HR diagram above the region boundary. Figure 1 also shows the boundaries of the fundamental-mode radial oscillations. The helium stars located in the HR diagram to the left of this boundary pulsate in overtones.

In Fig. 2, the normalized radial-displacement amplitude $\Delta R/R$ and the velocity amplitude ΔU at the

outer boundary of the model are plotted against the effective temperature T_{eff} for several sequences of model helium stars with a mass $M = 10\,M_\odot$. A comparison of Figs. 1 and 2 shows that the increase in T_{eff} in the lower part of the instability region at constant luminosity L is accompanied by a simultaneous decrease in the radial-displacement amplitude $\Delta R/R$ and the velocity amplitude ΔU of the outer boundary. This behavior is attributable to the existence of a high-temperature boundary of the pulsation-instability region that coincides with the boundary of the fundamental mode. As the luminosity rises, the instability boundary moves toward higher effective temperatures; the velocity amplitude even slightly increases with decreasing radial-displacement amplitude and reaches a maximum at $T_{\mathrm{eff}} \approx 7 \times 10^4$ K.

Such a different behavior of the plots of $\Delta R/R$ and ΔU against T_{eff} at high luminosities L stems from the fact that the instability-excitation zone moves closer to the photospheric level as the effective temperature rises. For pulsations in a particular overtone, it is necessary that the instability-excitation zone not overlap with the nodes of this overtone and be near one of its antinodes. Therefore, the decrease in the thickness of the ionization zone and its displacement toward the stellar surface are mainly responsible for the transition to pulsations in higher order overtones. In turn, the reduction in the pulsation period with increasing overtone order at a constant velocity amplitude ΔU causes the radial-displacement amplitude $\Delta R/R$ to decrease.

The transition from the fundamental mode to overtones can be clearly traced by examining the phase diagram of the fundamental Fourier harmonic in the complex (ξ_r, ξ_i) plane, where ξ_r and ξ_i are the real and imaginary parts of the normalized radial displacement amplitude $\Delta r/R$. Each Lagrangian layer of a pulsating model star is represented in the phase diagram by a point whose radius vector is proportional to the displacement amplitude, and the displacement phase is equal to the angle made by the radius vector with the horizontal axis: $\varphi = \tan^{-1}(\xi_i/\xi_r)$. Adiabatic standing-wave-type oscillations of the ℓth overtone are represented in the phase diagram by a straight line that crosses the coordinate origin ℓ times, while the phase diagram for adiabatic fundamental-mode ($\ell = 0$) oscillations is a straight line that emerges from the coordinate origin.

Figure 3 shows the phase diagrams of the fundamental Fourier harmonic of the radial displacement for three model helium stars with a mass $M = 10\,M_\odot$ and effective temperatures $T_{\mathrm{eff}} = 2.5 \times 10^4$, 5×10^4, and 7.5×10^4 K. The instability-excitation region is highlighted in each curve by a bold segment and the

Parameters of helium stars in the upper part of the pulsation instability region

$M/M_\odot$	$L/10^3 L_\odot$	$T_{\text{eff}}/10^3$ K	ΔU, km s^{-1}	$\Delta R/R$	Π, days	Q, days	$U_{\max}/v_e$	$\bar{R}/R_{\text{ph}}$
1	10	40	221	0.164	0.1159	0.03836	0.25	1.19
1	10	80	248	0.047	0.00629	0.01667	0.12	1.03
1	10	100	248	0.032	0.00265	0.01373	0.09	1.01
1	20	60	395	0.735	0.07124	0.04732	0.46	1.74
1	20	80	571	0.833	0.08245	0.1298	0.68	2.30
1	20	100	672	0.507	0.06084	0.1871	0.49	2.05
10	100	40	334	0.132	0.1387	0.02581	0.22	1.05
10	100	60	313	0.041	0.02280	0.01432	0.11	1.03
10	100	80	314	0.023	0.00746	0.01112	0.08	1.01
10	100	100	200	0.010	0.00270	0.00787	0.05	1.00
10	200	40	417	0.590	0.4350	0.04813	0.47	1.42
10	200	60	618	0.552	0.3873	0.1447	0.38	1.66
10	200	80	652	0.235	0.08834	0.07820	0.26	1.64
10	200	100	681	0.193	0.03371	0.05830	0.17	1.33

layer with maximum positive work per cycle $\oint P dV$, where P is the total pressure and V is the specific volume, is indicated by a filled circle. As we see from this figure, all of the phase diagrams exhibit appreciable variations in the phase φ along the curve, which is indicative of significantly nonadiabatic pulsations in the outer layers of a pulsating star. In particular, the phase decreases with increasing layer radius r up to the instability-excitation zone and increases above this zone. This behavior suggests the presence of two traveling waves that emerge at the time of maximum compression in the instability-excitation zone and propagate in opposite directions.

As we see from Fig. 3, the radial pulsations at $M = 10\,M_\odot$, $L = 10^5 L_\odot$, and $T_{\text{eff}} = 2.5 \times 10^4$ result from the instability of the fundamental mode. From the inner boundary of the model to the layer of maximum positive work, the phase of the fundamental Fourier harmonic changes by less than 0.05 radians, and the nonadiabaticity effects show up mostly in the outer layers of the stellar envelope. The oscillations at $T_{\text{eff}} = 5 \times 10^4$ are in the first overtone, because the instability excitation zone is above the overtone node. Because of significant nonadiabaticity, the radial-displacement amplitude has only a shallow minimum at the node and the phase continuously changes with increasing radius r. Therefore, the overtone node in the phase diagram can be determined

by the segment of the curve closest to the coordinate origin.

At $T_{\text{eff}} = 7.5 \times 10^4$ K, the pulsations result from the instability of the overtones of orders $k \gtrsim 6$. The phase diagram shown in Fig. 3 for this model represents the variations only in the outer layers with a radius $r \gtrsim 0.9R$, while in deeper layers the radial dependence of the amplitude has several minima and the phase of the pulsation wave changes several times by $\pi/2$ radians.

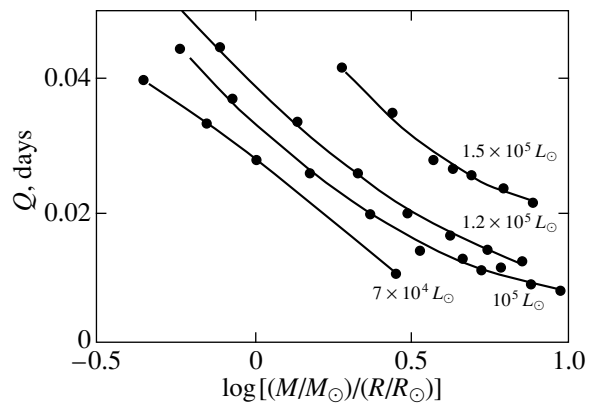The continuous change in the phase with in-

Fig. 4. The pulsation constant Q versus the logarithm of the mass-to-radius ratio for helium stars with a mass $M = 10\,M_\odot$. The luminosities L are given near the curves. The individual models are represented by the filled circles.
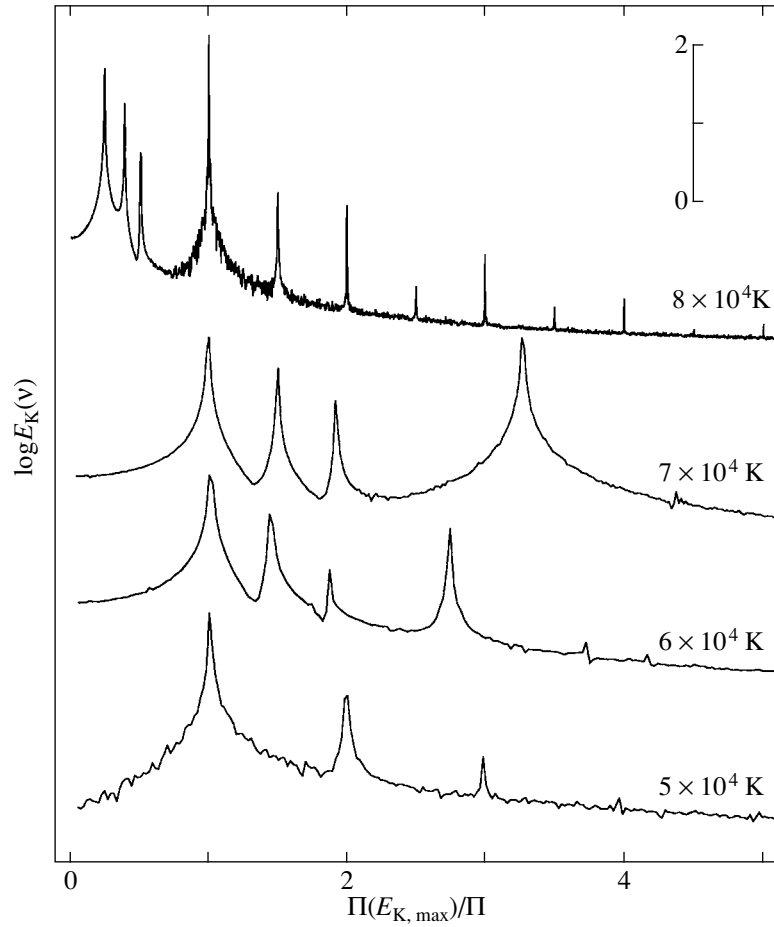
Fig. 5. Fourier spectra of the kinetic energy for model helium stars with a mass $M = 10\,M_\odot$ and a luminosity $L = 10^5 L_\odot$. The effective temperatures are given near the curves.

creasing distance from the stellar center and the fact that the phase difference between the inner and outer boundaries of the model is not a multiple of $\pi/2$ radians suggest that the pulsations cannot be described in terms of standing waves. For this reason, the transition to oscillations in higher order modes is accompanied by a continuous change in the pulsation constant Q. This peculiarity clearly distinguishes helium stars from classical pulsating δ Cep and RR Lyr variables, in which the transition from fundamental-mode oscillations to first-overtone oscillations is accompanied by an abrupt change in the pulsation constant. The pulsation constant Q is plotted against $f = \log[(M/M_\odot)/(R/R_\odot)]$ for $M = 10\,M_\odot$ and several luminosities L in Fig. 4; the pulsation periods Π and the values of Q for helium stars with luminosity-to-mass ratios $L/M = 10^3$ and $2 \times 10^3 L_\odot/M_\odot$ are given in the table.

The increase in the effective temperature of the star at $T_{\rm eff} \gtrsim 5 \times 10^4$ K and the transition to pulsations in higher-order overtones eventually cause the pulsations to become irregular. To elucidate the reason why

the pulsations become nonperiodic, we considered the spectral density of the kinetic energy of the stellar envelope

$$E_{\rm K}(\nu) = \frac{1}{2} \int_0^M \mathcal{U}(\nu)^2 dm, \tag{1}$$

where $dm = 4\pi r^2 \rho(r) dr$ is the element of mass, $\rho(r)$ is the gas density in a layer of radius r, and

$$\mathcal{U}(\nu) = \int_{-\infty}^{\infty} U(m, t) e^{2\pi i \nu t} dt \tag{2}$$

is the Fourier transform of the velocity $U(m, t)$ of the Lagrangian layer with the mass coordinate m. Figure 5 shows a sequence of $E_{\rm K}(\nu)$ plots for $M = 10\,M_\odot$, $L = 10^5\,L_\odot$ and several values of $T_{\rm eff}$. For the convenience of graphical representation, we use the dimensionless frequency $\nu/\nu_1 = \Pi_1/\Pi$, where Π_1 is the period that corresponds to the principal maximum in the spectrum $E_{\rm K}(\nu)$, as an independent variable.
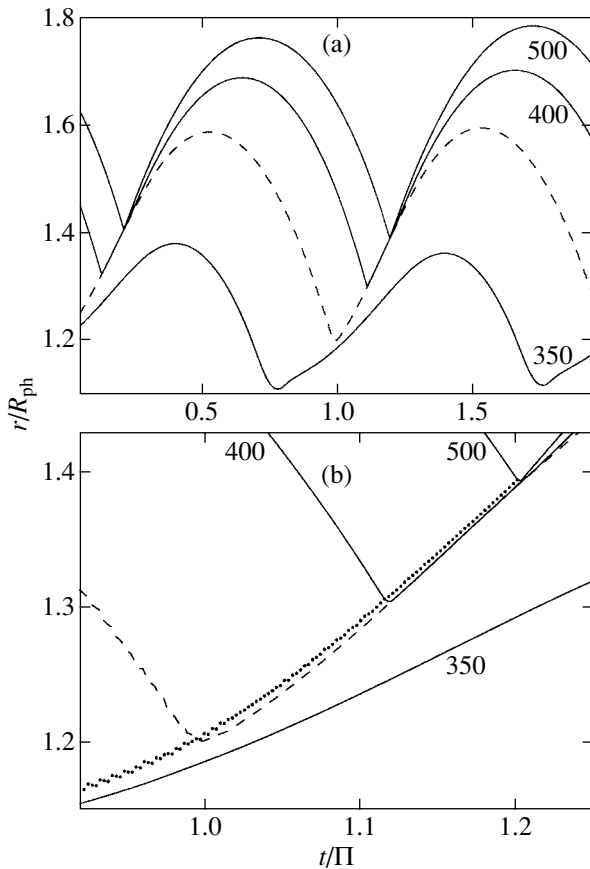
Fig. 6. (a) The radii of the outer Lagrangian layers (solid lines) and the photospheric radius (dashed line) versus the pulsation phase t/Π for a model with $M = 10\,M_\odot$, $L = 2 \times 10^5 L_\odot$, and $T_{\mathrm{eff}} = 8 \times 10^4$ K. The Lagrangian-layer numbers are given near the curves. The $j = 500$ layer is the outer boundary of the model. (b) The propagation of the velocity jump (filled circles) near the minimum of the photospheric radius.
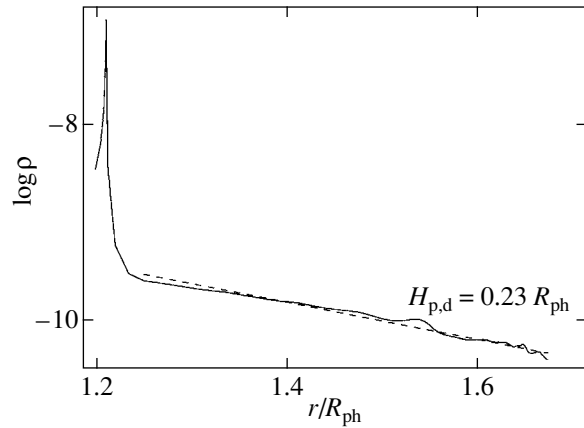

Fig. 7. The gas density ρ of the outer layers versus the distance from the center r for a model helium star with $M = 10\,M_\odot$, $L = 2 \times 10^5 L_\odot$, and $T_{\mathrm{eff}} = 8 \times 10^4$ K at $t/\Pi = 0$. The dashed line indicates the segment of the curve within which the gas-density distribution is described by the dynamic scale height $H_{\mathrm{p,d}}$.

For strictly periodic pulsations to exist, the maxima of the kinetic-energy spectral density $E_{\mathrm{K}}(\nu)$ must have a power-law distribution. For the models shown in Fig. 5, this condition is satisfied only at $T_{\mathrm{eff}} = 5 \times 10^4$ K. At higher values of T_{eff}, two or even more overtones are simultaneously excited (the models with $T_{\mathrm{eff}} \geq 7 \times 10^4$ K in Fig. 5), which causes the pulsations to become nonperiodic.

STELLAR ATMOSPHERES UNDER PULSATION CONDITIONS

Because of significant nonadiabaticity, the pulsations of the outer layers of the stellar envelope lag behind those of the deeper layers. The lag is illustrated by Fig. 6a, in which the photospheric radius and the radii of three Lagrangian layers are plotted against t/Π for a model helium star with $M = 10\,M_\odot$, $L = 2 \times 10^5\,L_\odot$, and $T_{\mathrm{eff}} = 8 \times 10^4$ K. The $j = 500$

layer is the outer boundary of the model, and the time t is measured from the time the photospheric radius is at a minimum. The radial dependence of the pulsation phase suggests the presence of a traveling wave that emerges in each oscillation cycle at the time of maximum compression in the instability-excitation zone. The trajectory of the traveling wave is shown in Fig. 6b for the same model on an enlarged scale. Note that the discontinuous pattern of variations in the coordinates of the traveling wave stems from the fact that the hydrodynamic model is discrete. Above the photospheric level, the traveling wave transforms into a shock wave.

The propagation of the shock wave through the stellar atmosphere during each pulsation cycle produces two important effects. First, the shock amplitude in a medium with a decreasing gas density increases with distance from the stellar center, so the expansion velocity of the outermost layers becomes comparable to the local escape velocity v_{e}. The maximum ratios of the expansion velocity for the outer boundary of the model to the local escape velocity, $U_{\max}/v_{\mathrm{e}}$, are given in the table. As follows from the data in this table, both an increase in L/M and a decrease in T_{eff}, i.e., an enhancement of the pulsation nonadiabaticity and a decrease of the gravity in the outer layers of the pulsating star, favor the shock gas-acceleration mechanism.

Second, an important consequence of periodic shock waves is the fact that after the completion of the stage of amplitude growth, the mean radii of the outer layers appreciably exceed the values that correspond to the initial hydrostatic and thermal equilibrium. This is clearly seen from the plots shown in Fig. 6, where all quantities are given in units of the equilibrium

photospheric radius R_{ph}. For example, the radius of the outer boundary in an equilibrium is larger than the photospheric radius by less than one percent, while the mean radius of the outer boundary for the dynamic model $\bar{R}$ can be more than twice as large as R_{ph}. The values of $\bar{R}/R_{ph}$ are given in the last column of the table.

The increase in the mean radius of the outer layers is accompanied by significant changes in the distribution of the gas density ρ in the stellar atmosphere. In Fig. 7, the gas density ρ is plotted against the distance from the stellar center r at the time $t/\Pi = 0$ that immediately precedes the onset of photospheric expansion. As we see from this plot, the gas-density distribution in the stellar atmosphere can be described in terms of the dynamic scale height $H_{p,d}$, whose value is more than an order of magnitude larger than the static scale height $H_p = P/(\rho g)$. For example, for the model with $M = 10\,M_\odot$, $L = 2 \times 10^5\,L_\odot$, and $T_{eff} = 8 \times 10^4$ K shown in Fig. 7, the ratio of the static scale height to the equilibrium photospheric radius is $H_p/R_{ph} \approx 0.006$, while the dynamic scale height is $H_{p,d} \approx 0.23 R_{ph}$.

CONCLUSIONS

It follows from the results of our hydrodynamic calculations presented above that for a heavy-element mass fraction $Z \geq 0.01$ and a luminosity-to-mass ratio $L/M \gtrsim 10^3 L_\odot/M_\odot$, the instability of helium stars against radial pulsations shows up in a wide range of effective temperatures. The general trends characteristic of helium stars with masses $1\,M_\odot \leq M \leq 10\,M_\odot$ and a heavy-element mass fraction $Z = 0.02$ are as follows. Pulsations emerge near the lower boundary of the instability region $(L/M \sim 10^3 L_\odot/M_\odot)$ at $T_{eff} \lesssim 3 \times 10^4$ K; they are attributable to the instability of the fundamental mode. As the luminosity increases, the temperature boundary of the instability region moves into the left part of the HR diagram and goes beyond $T_{eff} \approx 10^5$ K at $L/M \gtrsim 10^4 L_\odot/M_\odot$. In high-luminosity helium stars, fundamental-mode pulsations emerge at $T_{eff} \lesssim 4 \times 10^4$ K, while at higher effective temperatures the pulsations result from the instability of the overtones. The pulsations become nonperiodic both because of the nonlinearity at $\Delta R/R \sim 1$ and at high effective temperatures. In the latter case, the periodicity is disrupted, because two or more overtones with nonmultiple periods are simultaneously excited. This peculiarity can cause difficulty in identifying the frequencies of the observed light variations with pulsation modes.

The velocity amplitude of the outer layers ΔU increases with luminosity, and the dynamic-instability boundary that corresponds to the expansion of the outer layers at the velocity equal to the escape velocity for $T_{eff} \approx 7 \times 10^4$ K is roughly determined by a ratio $L/M \sim 3 \times 10^4 L_\odot/M_\odot$. However, it should be emphasized that even at $U_{max} < v_e$, radial stellar pulsations cause significant changes in the gas-density distribution in the stellar atmosphere. These changes can be described in terms of the dynamic scale height, which is more than an order of magnitude larger than the equilibrium scale height.

We chose data on 25 Wolf−Rayet stars with masses $M \lesssim 10\,M_\odot$ from Hamann et al. (1993, 1995), Hamann and Koesterke (1998), Nugis and Lamers (2000), Heger and Langer (1996), and de Marco and Schmutz (1999) and compared their positions in the HR diagram with the results of our calculations. All of these stars proved to be within the pulsation-instability region. Thus, taking into account the existing uncertainties in the estimated masses, luminosities, and effective temperatures of Wolf−Rayet stars, we, nevertheless, can assume with confidence that the significant increase in the gas density needed to reconcile the radiative stellar wind model with the observed mass-loss rates can be produced by the periodic shock waves that accompany radial stellar pulsations.

In our next study, we will compute the radial pulsations of helium stars with masses $M > 10\,M_\odot$ and pay much more attention to the dynamics of the outer atmospheric layers in pulsating stars.

ACKNOWLEDGMENTS

REFERENCES

1. P. Bratschi and A. Blecha, Astron. Astrophys. **313**, 537 (1996).
2. A. N. Cox and J. H. Cahn, Astrophys. J. **326**, 804 (1988).
3. Yu. A. Fadeyev, Astrophys. Space Sci. **213**, 75 (1994).
4. Yu. A. Fadeyev, Astron. Zh. **75**, 750 (1998) [Astron. Rep. **42**, 662 (1998)].
5. Yu. A. Fadeyev, Astron. Zh. **77**, 214 (2000) [Astron. Rep. **44**, 185 (2000)].
6. Yu. A. Fadeyev and M. F. Novikova, Astron. Zh. **80**, 75 (2003) [Astron. Rep. **47**, 68 (2003)].
7. A. Gautschy, Astron. Astrophys. **302**, 401 (1995).
8. W. Glatzel and S. Mehren, Mon. Not. R. Astron. Soc. **282**, 1470 (1996).
9. W. Glatzel, M. Kiriakidis, and K. J. Fricke, Mon. Not. R. Astron. Soc. **262**, L7 (1993).
10. W. Glatzel, M. Kiriakidis, S. Chernigovskij, and K. J. Fricke, Mon. Not. R. Astron. Soc. **303**, 116 (1999).

11. W.-R. Hamann and L. Koesterke, Astron. Astrophys. **333**, 251 (1998).
12. W.-R. Hamann, L. Koesterke, and U. Wessolowski, Astron. Astrophys. **274**, 397 (1993).
13. W.-R. Hamann, L. Koesterke, and U. Wessolowski, Astron. Astrophys. **299**, 151 (1995).
14. A. Hegera and N. Langer, Astron. Astrophys. **315**, 421 (1996).
15. J. E. Herald, R. E. Schulte-Ladbeck, P. R. J. Eenens, and P. Morris, Astrophys. J., Suppl. Ser. **126**, 469 (2000).
16. C. A. Iglesias and F. J. Rogers, Astrophys. J. **443**, 460 (1995).
17. C. A. Iglesias and F. J. Rogers, Astrophys. J. **464**, 943 (1996).
18. M. Kiriakidis, W. Glatzel, and K. J. Fricke, Mon. Not. R. Astron. Soc. **281**, 406 (1996).
19. A. Maeder, Astron. Astrophys. **147**, 300 (1985).
20. S. V. Marchenko, A. F. J. Moffat, T. Eversberg, *et al.*, Mon. Not. R. Astron. Soc. **294**, 642 (1998).
21. O. de Marco and W. Schmutz, Astron. Astrophys. **345**, 163 (1999).
22. P. Monderen, C. W. H. de Loore, K. A. van der Hucht, and A. M. van Genderen, Astron. Astrophys. **195**, 179 (1988).
23. T. Morel, L. N. Georgiev, Y. Grosdidier, *et al.*, Astron. Astrophys. **349**, 457 (1999).
24. T. Nugis and H. J. G. L. M. Lamers, Astron. Astrophys. **360**, 227 (2000).
25. F. J. Rogers, F. J. Swenson, and C. A. Iglesias, Astrophys. J. **456**, 902 (1996).
26. G. Schaller, in *Proceedings of the Conference on Confrontation between Stellar Pulsations and Evolution, Bologna, Italy, 1990,* ASP Conf. Ser. **11**, 304 (1990).
27. M. J. Seaton, Y. Yan, D. Mihalas, and A. K. Pradhan, Mon. Not. R. Astron. Soc. **266**, 805 (1994).
28. P. M. Veen, A. M. van Genderen, K. A. van der Hucht, *et al.*, Astron. Astrophys. **385**, 585 (2002a).
29. P. M. Veen, A. M. van Genderen, P. A. Crowther, and K. A. van der Hucht, Astron. Astrophys. **385**, 600 (2002b).
30. P. M. Veen, A. M. van Genderen, and K. A. van der Hucht, Astron. Astrophys. **385**, 619 (2002c).

Translated by V. Astakhov

Shock Acceleration of Solar Cosmic Rays

E. G. Berezhko and S. N. Taneev[*]

Institute of Cosmophysical Research and Aeronomy, Siberian Branch, Russian Academy of Sciences,
pr. Lenina 31, Yakutsk, 677891 Russia
Received February 11, 2003

Abstract—We investigated the acceleration of solar cosmic rays (SCRs) by the shock waves produced by coronal mass ejections. We performed detailed numerical calculations of the SCR spectra produced during the shock propagation in the solar corona in terms of a model based on the diffusive transport equation using a realistic set of physical parameters for the corona. The resulting SCR energy spectrum $N(\varepsilon) \propto \varepsilon^{-\gamma} \exp\left[-\left(\varepsilon/\varepsilon_{\max}\right)^{\alpha}\right]$ is shown to include a power-law portion with an index $\gamma \simeq 2$ that ends with an exponential tail with $\alpha \simeq 2.5 - \beta$, where β is the spectral index of the background Alfvén turbulence. The maximum SCR energy lies within the range $\varepsilon_{\max} = 1-300$ MeV, depending on the shock velocity. Because of the steep spectrum of the SCRs, their backreaction on the shock structure is negligible. The decrease in the Alfvén Mach number of the shock due to the increase in the Alfvén velocity with heliocentric distance r causes the efficient SCR acceleration to terminate when the shock reaches a distance of $r = 2-3R_{\odot}$. Since the diffusive SCR propagation in this case is faster than the shock expansion, SCR particles intensively escape from the shock vicinity. A comparison of the calculated SCR fluxes expected near the Earth's orbit with available experimental data indicates that the theory satisfactorily explains all of the main observed features. © 2003 MAIK "Nauka/Interperiodica".

Key words: *solar corona, shock waves, particle diffusive shock acceleration, solar cosmic rays, and Alfvén waves.*

INTRODUCTION

The so-called gradual enhancements (or events) of solar energetic particles (which are called solar cosmic rays (SCRs) below) are currently believed to be generated in the solar corona at the shock fronts produced by coronal mass ejections (CMEs) (see, e.g., the review articles by Reames (1996, 1999, 2000); the book by Miroshnichenko (2001) and references therein). While the relationship between gradual SCR enhancements and CMEs directly follows from measurements (Kahler *et al.* 1978; see also the review article by Reames (2000) and references therein), a detailed theoretical analysis is required to determine the SCR acceleration mechanism.

Diffusive shock acceleration, which was originally established by Krymskii (1977) and Axford *et al.* (1977) (see also the review article by Berezhko and Krymskii (1988) and the monograph by Berezhko *et al.* (1988)), is considered as the most appropriate SCR acceleration mechanism.

The first attempts to apply the theory of diffusive shock acceleration to the solar coronal conditions (Ellison and Ramaty 1985; Lee and Ryan 1986) showed that SCR spectra with the required properties

could be generated in principle. At the same time, these studies contain several significant simplifications that restrict their predictive possibilities. One of them, the plane-wave approximation (Ellison and Ramaty 1985; Zank *et al.* 2000), disregards the finite shock size and its time dependence. This approximation is valid for the bulk of the accelerated particles that form a power-law spectrum near the shock front at each current instant of time. At the same time, it breaks down near or above the maximum energy where the accelerated-particle spectrum undergoes an exponential cutoff. In many cases, the finiteness of the shock size is the major physical factor that determines the maximum energy $\varepsilon_{\max}$ and the particular features of the spectral shape at $\varepsilon \gtrsim \varepsilon_{\max}$.

The exponential cutoff in the accelerated-particle spectrum takes place in the range of energies at which the diffusion length $l = \kappa(\varepsilon)/V_{\mathrm{S}}$ reaches $l_{\max} = l(\varepsilon_{\max}) \simeq 0.1 R_{\mathrm{S}}$. Here, R_{S} and V_{S} are the shock size and velocity, respectively, and $\kappa(\varepsilon)$ is the particle diffusion coefficient. Hence, a realistic diffusion coefficient $\kappa(\varepsilon)$ should be used to properly describe the particle acceleration in the range of maximum energies. Such important SCR characteristics as their maximum energy $\varepsilon_{\max}$ and spectral shape at $\varepsilon \gtrsim \varepsilon_{\max}$ cannot be reproduced using a model with

[*]E-mail: `taneev@ikfia.ysn.ru`

an ε-independent diffusion coefficient κ (Lee and Ryan 1986).

The particular features of the exponential portion of the spectrum are of special interest, because the highest-energy SCRs recorded during the most intense events by ground-based detectors (neutron monitors and meson telescopes) belong precisely to the exponential part of the SCR spectrum.

An important property of the SCRs observed far from the Sun, $r \gg R_\odot$, is the fact that particles with energy $\varepsilon \gtrsim 1$ MeV come to the observer much earlier than the shock. Since we deal with the same shock that produced SCRs at the onset of its propagation from the Sun, the acceleration efficiency significantly changes or, more specifically, decreases as the shock propagates. As we show below, the acceleration efficiency decreases within heliocentric distances $r < 4R_\odot$ because the shock weakens (i.e., its Alfvén Mach number M_a decreases) and the diffusion coefficient $\kappa(\varepsilon, r)$ increases with distance r. For this reason, most of the SCRs are produced at $r < 3R_\odot$. Subsequently, the diffusive SCR propagation prevails over the shock-expansion velocity. Therefore, a remote observer records the SCR arrival much earlier than the shock arrival.

Here, our main goal is to theoretically calculate the shock-generated spectra in the solar corona and to compare the results of our calculations with experimental data. In contrast to previous studies, our calculations are based on realistic parameters of the solar corona and shock velocities as well as on an SCR diffusion coefficient that is consistent with current views of the Alfvén turbulence in the solar corona. As we show below, the theory satisfactorily explains all of the main features of the SCR streams in gradual events observed near the Earth's orbit.

PARTICLE ACCELERATION

The problem under consideration was formulated in general terms previously (Berezhko *et al.* 1998, 2001a, 2001b). Here, we describe the main content of the model in more detail.

The front of a CME-driven shock probably has a complex, nonspherical shape. It may be assumed that the particle acceleration is most efficient at the shock forefront with the highest velocity V_S and that the magnetic-field lines **B** make a small angle θ with the normal to the shock front. Therefore, we will represent the front of a traveling shock as a segment of a spherical surface of radius $R_S(t)$ that increases with time at a constant velocity $V_S = dR_S/dt$ and that corresponds to the solid angle Ω_S. Within this segment, the shock is assumed to be purely parallel.

In the range of heliocentric distances $r < 10R_\odot$ under consideration, we assume that the magnetic field **B** is directed radially as the flow velocity **w**. Since the half-width of the characteristic cross section $L_\perp$ of the forefront (i.e., the acceleration region) is fairly large ($L_\perp \sim R_S$) and since fast particles are strongly magnetized ($\kappa_\parallel \gg \kappa_\perp$), the approximation of spherical symmetry in our case implies that all physical quantities are functions of only one spatial variable—the heliocentric distance r. In this case, the diffusive transport equation for the particle-distribution function $f(r, p, t)$ is

$$\frac{\partial f}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(\kappa_\parallel r^2 \frac{\partial f}{\partial r}\right) - w' \frac{\partial f}{\partial r} \qquad (1)$$

$$+ \frac{p}{3r^2}\frac{\partial(w'r^2)}{\partial r}\frac{\partial f}{\partial p} - \frac{f}{\tau_\perp},$$

where $\kappa_\parallel$ ($\kappa_\perp$) is the parallel (perpendicular) (with respect to the magnetic field **B**) particle diffusion coefficient and p is the particle momentum. Since Alfvén waves in the solar corona travel mostly away from the Sun (see, e.g., Velli and Pruneti (1997) and references therein), the scattering centers at $r > R_S$ have the velocity $w' = w + c_a$, where $c_a = B/\sqrt{4\pi\rho}$ is the Alfvén velocity and ρ is the density of the medium. In the downstream region ($r < R_S$), the propagation of Alfvén waves is largely isotropized; therefore, $w' = w$.

The third term on the right-hand side of Eq. (1) describes the adiabatic particle deceleration in an expanding flow. This is one of the factors that limits the accelerated-particle spectrum at high energies.

The last term in Eq. (1) effectively describes the particle escape from the acceleration region through perpendicular diffusion with the time scale $\tau_\perp = L_\perp^2/\kappa_\perp$. The actual values of the diffusion coefficient $\kappa_\perp$ are such that the term $f/\tau_\perp$ has only a weak effect on the particle acceleration. As in the case of particle acceleration by interplanetary shocks (i.e., at distances $r > 0.1$ AU) (Berezhko *et al.* 1998), we assume that $L_\perp = 0.6R_S$, which corresponds to $\Omega_S = 1.26$ sr.

Note that the angle Ω_S affects only the total number of shock-produced SCRs (which is directly proportional to Ω_S) and has absolutely no effect on their distribution within the cone angle Ω_S.

We disregard the shock modification by the back-reaction of the accelerated SCRs, because their pressure, as we show below, is much lower than the dynamic pressure ρV_S^2. Therefore, the shock front is treated as a discontinuity at which the velocity of the medium relative to the shock front $u = V_S - w$ abruptly changes from u_1 at $r = R_S + 0$ to $u_2 = u_1/\sigma$ at $r = R_S - 0$, where

$$\sigma = 4 / \left(1 + 3/M_1^2\right) \qquad (2)$$

is the shock compression ratio, $M = u/c_s$ is the Mach number, $c_s = \sqrt{\gamma_g k_B T/m}$ is the speed of sound, T is the temperature, k_B is the Boltzmann constant, m is the proton mass, and the polytropic index of the plasma was taken to be $\gamma_g = 5/3$; the subscript "1" ("2") refers to points just ahead of (behind) the shock front.

The distribution function at the shock front ($r = R_S$) satisfies the condition

$$\frac{u_1' - u_2}{3} p \frac{\partial f}{\partial p} = \left(\kappa_\parallel \frac{\partial f}{\partial r} \right)_1 - \left(\kappa_\parallel \frac{\partial f}{\partial r} \right)_2 + Q_0, \quad (3)$$

in which $u_1' = u_1 - c_a$ and

$$Q_0 = u_1 \frac{N_{inj}}{4\pi p_{inj}^2} \delta(p - p_{inj}) \quad (4)$$

is the source concentrated at the shock front that injects a fraction $\eta = N_{inj}/N_{g1}$ of the gas particle number density $N_{g1} = N_g(r = R_S + 0)$ into acceleration. For simplicity, we restrict our analysis to protons—the main type of ions in the coronal plasma, which is assumed to be a purely hydrogen plasma. Therefore, the coronal density ρ and the proton number density N_g are related by $\rho = mN_g$, where m is the proton mass.

Since the diffusive particle mobility increases with energy, the most energetic particles of the medium after their shock heating can again cross the shock front in the opposite direction, which implies their injection into acceleration. Theoretical analysis (Malkov and Völk 1995), numerical simulations of collisionless shocks (Scholer et al. 1992), and measurements at the Earth's bow shock (Trattner et al. 1994) strongly suggest that there is an efficient injection of superthermal particles: apart from the thermal (Maxwellian) peak, the ion energy distribution at the front of a quasi-parallel shock has a high-energy power-law tail that is formed by diffusive shock acceleration. These particles, which we arbitrarily call cosmic rays, are the subject of our study, as applied to the shock propagation in the solar corona.

Since there is no well-developed theory of the injection mechanism (or, to be more precise, theory of the strong shock transition), η is a free parameter in our model. Based on the analysis of experimental data (Trattner et al. 1994) and numerical simulations of collisionless shocks (Scholer et al. 1992; Trattner and Scholer 1993), we can only specify the possible range of values for this parameter, $\eta = 10^{-3}-10^{-2}$.

At the same time, we do not rule out a situation where the actual injection rate can be much lower ($\eta \ll 10^{-3}$) than that for a quasi-parallel shock, because a significant part of the shock front at each current instant of time is quasi-perpendicular. The reason is that the rate of particle injection into acceleration rapidly decreases with increasing angle θ for $\theta \gtrsim 45°$ (Ellison et al. 1995; Malkov and Völk 1995). The quasi-perpendicular parts of the front can prevail either because of the corresponding large-scale magnetic-field structure or in the presence of large-amplitude background turbulence. It may well be that both these factors are important in the solar corona.

The choice of the injected-particle momentum p_{inj} that, by its meaning, separates the slow (thermal) and fast (accelerated) particles is, to some extent, arbitrary. In fact, it is limited only by the validity condition for the entire range $p \geq p_{inj}$ of the diffusion approximation based on Eq. (1). Therefore, we assume, as usual, that $p_{inj} = \lambda m c_{s2}$, where $\lambda > 1$ (see, e.g., Berezhko et al. 1996) and $c_{s2} = u_1 \sqrt{\gamma_g(\sigma - 1) + \sigma/M_1^2}/\sigma$ is the speed of sound downstream of the shock front. In our calculations, we used $\lambda = 4$.

Since the shock front is the only source where particles are injected into acceleration, the problem should be solved under the following initial and boundary conditions:

$$f(r, p, t_0) = 0, \qquad f(r = \infty, p, t) = 0, \quad (5)$$

which imply the absence of background particles in the solar wind with energies in the range under consideration.

As in previous studies, we assume that the medium in the downstream region $r < R_S$ is perturbed much more strongly than it is in the upstream region ($r > R_S$), which ensures that $\kappa_{\parallel 2} \ll \kappa_{\parallel 1}$.

This assumption allows the second term on the left-hand side of Eq. (3) to be disregarded. As a result, the solution of the problem does not depend on any singularities of the region $r < R_S$. In this case, the particle distribution function in this region can easily be determined from Eq. (1) if we disregard the diffusion term, which is justified by the adopted condition, for any given velocity field w:

$$f(r, p, t) = f_R(p_R, t_R), \quad (6)$$

where $f_R(p, t) = f(r = R_S, p, t)$ is the particle-distribution function at the shock front; $p_R = ap$;

$$a = \exp \left[\frac{1}{3} \int_{t_R}^{t} \nabla \mathbf{w}(s, t') dt' \right] \quad (7)$$

is the factor that describes the adiabatic particle deceleration in the downstream region $r < R_S$; and $s(t)$ is the solution of the equation

$$ds/dt = w(s, t) \quad (8)$$

with the boundary conditions

$$s(t_R) = R_S(t_R), \qquad s(t) = r. \qquad (9)$$

To determine the distribution function of the downstream accelerated particles requires specifying the velocity field $w(r)$ in this region. Since there is no reliable information on the distribution $w(r)$ in the literature, we take it here in the form

$$w = w_2 r/R_S, \qquad (10)$$

which corresponds to a constant $\nabla \mathbf{w} = 3w_2/R_S$ at $r < R_S$.

The diffusion coefficients in Eq. (1) are defined by the relations (Lee 1982, 1983)

$$\kappa_\parallel = \frac{v^2 B^2}{32\pi^2 \omega_B E\left(k = \rho_B^{-1}\right)}, \qquad \kappa_\parallel \kappa_\perp = \frac{\rho_B^2 v^2}{3}, \qquad (11)$$

where v is the particle velocity, $\rho_B = v/\omega_B$ is the gyroradius, $\omega_B = eB/mc$ is the gyrofrequency, e is the elementary charge, c is the speed of light, and $E(k) = d(\delta B^2/8\pi)/d\ln k$ is the energy density of the Alfvén waves. Particles are scattered due to their interaction only with the waves whose wave number k is equal to the inverse particle gyroradius ρ_B.

PARAMETERS OF THE SOLAR CORONA

For the radial proton-density distribution in the low-latitude corona, we use a semiempirical model of Sittler and Guhathakurta (1999):

$$N_g(r) = N_0 \left[a_1 e^{a_2 z} z^2 (1 + a_3 z + a_4 z^2 + a_5 z^3)\right], \qquad (12)$$

where $N_0 = 10^8$ cm^{-3}, $a_1 = 3.2565 \times 10^{-3}$, $a_2 = 3.6728$, $a_3 = 4.8947$, $a_4 = 7.6123$, $a_5 = 5.9868$, and $z = R_\odot/r$. For simplicity, we ignore all types of ions except protons. Therefore, the density of the medium in our case is $\rho = N_g m$.

The velocity of the medium can be determined from the continuity condition for the flow of matter under the assumption of spherical symmetry:

$$w(r) = w_e \left[N_e/N_g(r)\right] (r_e/r)^2, \qquad (13)$$

where $w_e = w(r_e) = 450$ km s^{-1} and $N_e = N_g(r_e) = 8$ cm^{-3} are, respectively, the velocity and density of the solar wind near the Earth's orbit $r = r_e = 1$ AU.

The temperature is taken to be $T = 1.2 \times 10^6$ K, a value that agrees with the data of Hundhousen (1972) in the distance range $r < 3R_\odot$ of interest. On the other hand, Reames (2000) obtained $T \simeq 2 \times 10^6$ K. This uncertainty in T is not important for our model, because, as we show below, the SCR acceleration efficiency is determined mainly by the Alfvén Mach
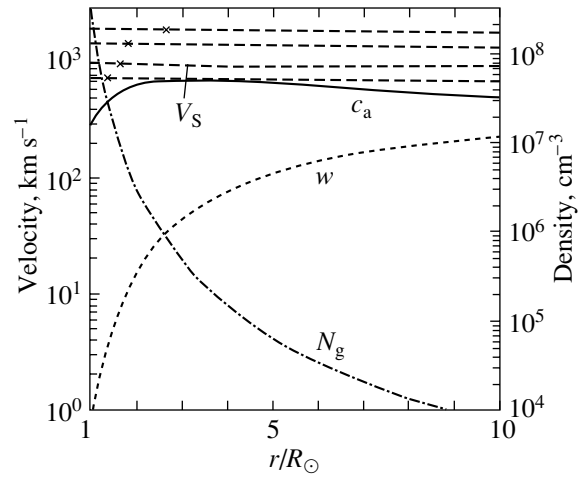
Fig. 1. The proton density of the medium N_g, its velocity w, and the Alfvén velocity c_a versus the heliocentric distance. The dashed lines correspond to the shock velocities $V_S = 750, 1000, 1500,$ and 2000 km s^{-1}; the crosses indicate the size of the efficient SCR acceleration region.

number $M_{a1} = u_1/c_a$ rather than by the sonic Mach number $M_1 = u_1/c_s$, which depends on the temperature of the medium.

The radial magnetic-field distribution $B(r)$ in the model of Sittler and Guhathakurta (1999) differs from the dependence $B \propto r^{-2}$ only at very small distances $r - R_\odot \ll R_\odot$. Therefore, as in the solar wind, we use the distribution

$$B(r) = B_\odot (R_\odot/r)^2, \qquad (14)$$

where we took $B_\odot = 2.3$ G for the solar surface magnetic-field strength, a value that is characteristic of the low heliolatitudes.

The radial profiles of proton density $N_g(r)$, velocity of the medium $w(r)$, and Alfvén velocity $c_a(r)$ in the distance range $1 \le r/R_\odot \le 10$ are shown in Fig. 1.

The least studied characteristic of the solar corona is the Alfvén-wave spectrum $E(k, r)$, which is important for the acceleration process. There are only a few general constraints that we use here. First, as in the solar wind (i.e., at distances $r > 10R_\odot$), the dependence of the Alfvén-wave energy content on the wave number k and heliocentric distance r is commonly represented as a power law

$$E(r, k) = E_0 (k/k_0)^{-\beta} (r/r_0)^{-\delta}, \qquad (15)$$

where it would be most natural to take $R_\odot$ as r_0 for the corona.

Since the wave–particle interaction is resonant in nature—waves with a wave number k scatter particles with a gyroradius $\rho_B = 1/k$—and since the particle-energy range of interest is bounded by several hundreds of MeV, we choose $k_0 = 1.4 \times 10^{-6}$ cm^{-1},

which corresponds to the gyroradius of a particle with momentum $p = 0.5mc$ in a magnetic field $B = B_\odot = 2.3$ G as the parameter k_0. Undoubtedly, there are also longer waves with $k < k_0$ in the solar corona. However, this circumstance is unimportant for subsequent analysis, because, in general, the long-wavelength part of the spectrum is accounted for by a minor fraction of the total energy content (Tu and Marsh 1995; You Qiu Hu et al. 1999).

The total energy content of the Alfvén waves with $k > k_0$ at the base of the solar corona is $E_w = E_0/\beta$. The Alfvén-wave fluxes into the solar corona at its base are considered one of the main energy sources that heat up the corona and produce the solar wind. Since the energy flux density required to produce the solar wind is known, $F_w \simeq 5 \times 10^5$ erg cm^{-2} s^{-1}, it may be considered as an upper limit for the Alfvén wave energy flux $F_w = 2c_a E_w$. Therefore, given that $c_a \simeq 300$ km s^{-1} at the base of the corona, the calculations presented below were performed for $E_0 \sim 10^{-3}$ erg cm^{-3}.

The parameters β and δ that determine the dependence of the Alfvén-wave energy content on the wave number and heliocentric distance, respectively, are well known only for large distances $r \gtrsim 0.5$ AU, i.e., for the solar-wind region where $\beta = 0.5$ and $\delta = 4$ (recall that we use the Alfvén-wave energy content per unit interval of the logarithm of the wave number as the spectrum $E(k)$).

At smaller heliocentric distances $3 \le r/R_\odot \le 16$, information on the Alfvén-wave spectrum can be extracted from coronal radio-sounding experiments. Analysis of such measurements indicates that the spectral index β slightly increases as one approaches the Sun and that the spatial dependence of the Alfvén turbulence at $r < 6R_\odot$ is determined by $\delta \simeq 8$ (Andreev et al. 1997). Since, according to the calculations presented below, the SCRs are efficiently accelerated at $r < 4R_\odot$, we used $\beta = 0.5-1.5$ and $\delta = 8$ in our calculations.

THE SCR PROPAGATION

Problem (1)–(11) to determine the distribution function of the accelerated particles can be solved numerically within the time interval from the onset of shock propagation t_0 to the time t_f the acceleration becomes inefficient. The shock size at t_f is $R_S = R_f = 1.3-3R_\odot$, depending on the shock velocity V_S.

Since the SCR fluxes are most frequently measured near the Earth's orbit, i.e., at distances much larger than R_f, a relationship should be established between the SCR spectrum formed during the acceleration and the expected SCR flux at succeeding times at large distances $r \gg R_\odot$.

A simplified formulation of the problem on the propagation of shock-accelerated SCRs appears as follows. The resulting differential momentum spectrum of the SCRs produced by the time t_f is defined by the expression:

$$N_f(p) = \Omega_S \int_0^\infty f(r, p, t_f) r^2 dr. \qquad (16)$$

The SCR propagation in interplanetary space is a difficult problem in itself. A detailed description of this process is based on the kinetic equation that adequately describes the pitch-angle SCR diffusion in the interplanetary magnetic field (Toptygin 1983). At the same time, if the SCR scattering mean free path λ is small, $\lambda \ll r$, then the diffusion approach (Kota et al. 1982) based on the diffusive transport equation (1) becomes acceptable. A further simplification is achieved if the condition $\kappa \gg rw$, which allows the effect of convective SCR transport in the solar wind to be disregarded, is also satisfied. At $r = 1$ AU, the two conditions are satisfied for $\lambda \sim 0.1$ AU, which are typical of the SCRs (Toptygin 1983; Miroshnichenko 2001). In this case, the SCR propagation in interplanetary space can be described by the simplified diffusive transport equation

$$\frac{\partial n}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(\kappa r^2 \frac{\partial n}{\partial r} \right) + Q \qquad (17)$$

for the differential number density $n(r, p, t) = 4\pi p^2 f(r, p, t)$, where the source term

$$Q = \frac{N_f(p)}{4\pi R_f^2} \delta(r - R_f) \delta(t - t_f) \qquad (18)$$

corresponds to the release of particles with the spectrum $N_f(p)$ at distance $r = R_f$ at time $t = t_f$. Disregarding R_f compared to $r \gg R_f$ under consideration, we can write the solution of Eq. (17) as (Krimigis 1965)

$$n(r, t) = \frac{N_f}{\Omega_s} \left[(2 - \mu)^{\frac{4+\mu}{2-\mu}} \Gamma\left(\frac{3}{2-\mu} \right) (a\Delta t)^{\frac{3}{2-\mu}} \right]^{-1}$$

$$\times \exp\left[-\frac{r^{2-\mu}}{a\Delta t (2-\mu)^2} \right],$$
(19)

where $\Gamma(x)$ is the gamma function, $\Delta t = t - t_f$, and the radial dependence of the SCR diffusion coefficient is taken to be a power law, $\kappa(p, r) = a(p) r^\mu$.

Since the magnetic field is nearly radial at $r < 1$ AU, we may assume that $\kappa \simeq \kappa_\parallel$. As in the acceleration region, the diffusion of the SCRs is determined by their interaction with Alfvén waves. Therefore, the diffusion coefficient is defined by expression (11),

in which the Alfvén-wave spectrum $E(k,r)$ corresponds to measurements near $r = 1$ AU. As follows from the analysis of the experimental data (Russell 1972) presented in Lee (1982), the spectrum near the Earth's orbit typical of quiet conditions can be represented as (15) with $\beta_e \approx 0.5$ and $\delta_e = 4$. If we take $r_{0e} = 1$ AU as r_0 in this expression and $k_{0e} = 1.1 \times 10^{-9}$ cm^{-1}, which corresponds to a particle energy of 0.1 MeV, then the spectrum amplitude will be $E_{0e} \sim 10^{-13}$ erg cm^{-3} (the parameters marked by the subscript "e" specify the Alfvén-wave spectrum in interplanetary space $10R_\odot < r \le 1$ AU).

For the specified radial dependence of the wave energy density $E \propto r^{-4}$ and with the adopted dependence $B \propto r^{-2}$, the SCR diffusion coefficient in the solar wind varies with distance as $\kappa \propto r \, (\mu = 1)$.

CALCULATIONS: SCR SPECTRA IN THE SOLAR CORONA

The results of our calculations presented below were obtained by numerically solving the problem formulated above.

According to observations (Kahler *et al.* 1999), only the fastest CMEs with velocities above 500 km s^{-1} are accompanied by detectable SCR fluxes. On the other hand, the number of CMEs whose velocities appreciably exceed 1200 km s^{-1} is small (Burkepile and Cyr 1993). Therefore, we restricted our analysis to CME-driven shocks with velocities $V_S = 750$, 1000, 1500, and 2000 km s^{-1}.

One might expect the velocity of the shock to change during its propagation in the solar corona. However, there are no reliable data on this point. We only know from observations that the velocity of the shock as it propagates up to $r = 1$ AU decreases by less than a factor of 2 (Reames 1999). Therefore, we restricted our analysis to the simplest case of shocks with a constant velocity. This assumption is also justified by the fact that, according to our calculations presented below, an efficient SCR acceleration ends when the shock size reaches $R_S = 2$–$3R_\odot$.

Figure 2 presents the results of our calculations for $\eta = 10^{-4}$, $V_S = 1000$ km s^{-1}, $E_0 = 1.2 \times 10^{-3}$ erg cm^{-3}, and $\delta = 8$. Figure 2a shows the differential (with respect to kinetic energy ε) intensity of the shock-accelerated protons,

$$J(\varepsilon) = p^2 f(r = R_S, p, t), \qquad (20)$$

and Fig. 2b shows the calculated differential (with respect to kinetic energy ε) number of protons produced by time t,

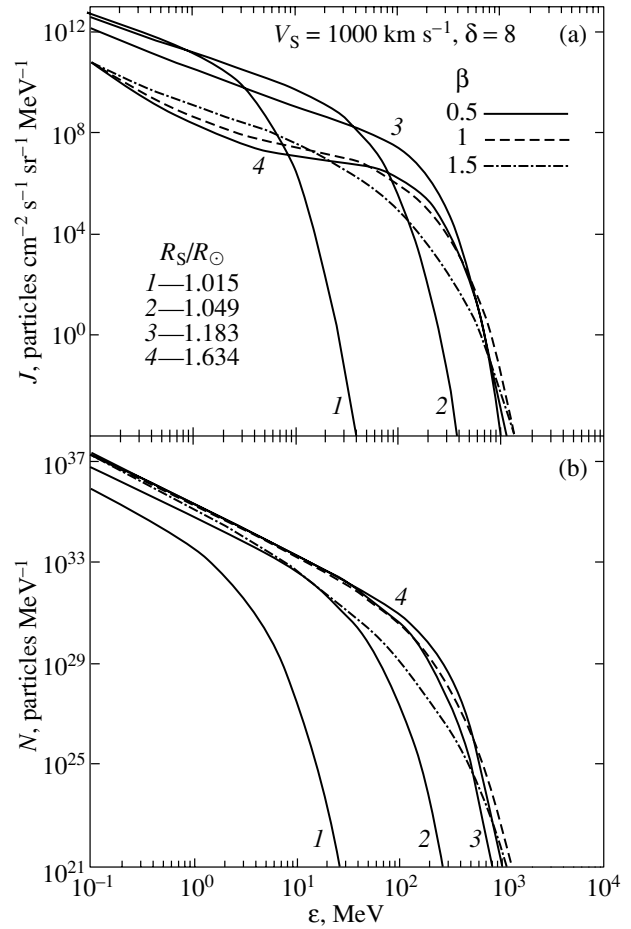$$N(\varepsilon) = \frac{4\pi p^2 \Omega_S}{v} \int_0^\infty f(r, p, t) r^2 dr, \qquad (21)$$

Fig. 2. (a) The differential intensity of the shock-accelerated protons and (b) their resulting spectrum versus the kinetic energy for a shock velocity $V_S = 1000$ km s^{-1}, injection rate $\eta = 10^{-4}$, Alfvén-wave energy density $E_0 = 1.2 \times 10^{-3}$ erg cm^{-3}, and index $\delta = 8$. For the spectral index $\beta = 0.5$ (solid lines), the spectra marked by numbers 1–4 correspond to shock sizes of $R_S/R_\odot = 1.015$, 1.049, 1.183, and 1.634, respectively. The spectra that correspond to $\beta = 1$ (dashed line) and $\beta = 1.5$ (dash–dotted line) are shown for $R_S/R_\odot = 1.634$.

where $\varepsilon = \sqrt{p^2 c^2 + m^2 c^4} - mc^2$.

For $\beta = 0.5$, the calculations are presented for four sequential times in order to show the contribution of different shock evolution phases to the resulting SCR spectrum $N_f(\varepsilon) = N(\varepsilon, t_f)$. For the other two cases, $\beta = 1$ and 1.5, the calculations are presented only for the time $t_f = 434$ s, when the shock size reaches $R_S(t_f) = 1.63R_\odot$ and the acceleration becomes inefficient.

The evolution of the accelerated-particle spectrum at the shock front consists in the formation of a

power-law spectrum (see, e.g., Berezhko *et al.* 1988)

$$f = \frac{qN_{\text{inj}}}{4\pi p_{\text{inj}}^3} \left(\frac{p}{p_{\text{inj}}}\right)^{-q} \qquad (22)$$

in the momentum range $p_{\text{inj}} \leq p < p_{\text{max}}$ with a spectral index

$$q = 3\sigma_{\text{eff}}/(\sigma_{\text{eff}} - 1), \qquad (23)$$

where

$$\sigma_{\text{eff}} = u_1'/u_2 = \sigma(1 - 1/M_{\text{a}1}), \qquad (24)$$

$\sigma = u_1/u_2$ is the shock compression ratio, and $M_{\text{a}1} = u_1/c_{\text{a}}(R_{\text{S}})$ is the current Alfvén Mach number. Figure 3 shows that the increase in Alfvén velocity c_{a} causes the spectral index q to increase as the shock propagates. The behavior of the maximum momentum $p_{\text{max}}(t)$ is similar to the case of an interplanetary shock (Berezhko *et al.* 1998). Initially, $p_{\text{max}}(t)$ monotonically increases with time, as in the case of a plane shock, in accordance with the equation

$$\frac{dp_{\text{max}}}{dt} = \frac{p_{\text{max}}}{\tau_{\text{a}}}, \qquad (25)$$

where $\tau_{\text{a}} = 3\kappa(R_{\text{S}})/[u_1'(u_1' - u_2)]$ is the acceleration time. At $p \gtrsim p_{\text{max}}$, the distribution function undergoes an exponential cutoff. We use the standard (in the theory of cosmic-ray acceleration) definition of the maximum momentum p_{max}: p_{max} is the momentum at which the accelerated-particle distribution function $f(p)$ is a factor of e smaller than the power-law spectrum $f \propto p^{-q}$, i.e., $f(p_{\text{inj}})(p_{\text{max}}/p_{\text{inj}})^{-q}/f(p_{\text{max}}) = e$.

The intensity $J(\varepsilon)$ shown in Fig. 2a behaves accordingly: a power-law portion $J(\varepsilon) \propto \varepsilon^{-\gamma}$ with an index $\gamma = (q-2)/2$ is formed in the range $\varepsilon_{\text{inj}} \leq \varepsilon < \varepsilon_{\text{max}} = p_{\text{max}}^2/2m$.

For the case shown in Fig. 2, the maximum energy $\varepsilon_{\text{max}}(t)$ virtually ceases to increase starting from the time that corresponds to $R_{\text{S}} \simeq 1.2R_{\odot}$, because the geometrical factor whose effect is related to the finite shock size becomes a more important factor that limits the accelerated-particle spectrum. In this case, the maximum momentum can be determined from the relation (Berezhko 1996)

$$g(p_{\text{max}}) \equiv \frac{R_{\text{S}}u_1}{\kappa_{\|}(p_{\text{max}})} = \frac{q(\delta - 2\beta - 2)}{2 - \beta}. \qquad (26)$$

The quantity $g(p)$ is called the modulation parameter. An efficient acceleration, as at a plane shock front when the effect of the finite shock size may be ignored, takes place in the range of momenta for which $g(p) \gg g(p_{\text{max}})$.

Formula (26) accurately defines the maximum momentum for a strong shock when the index $q \simeq 4$
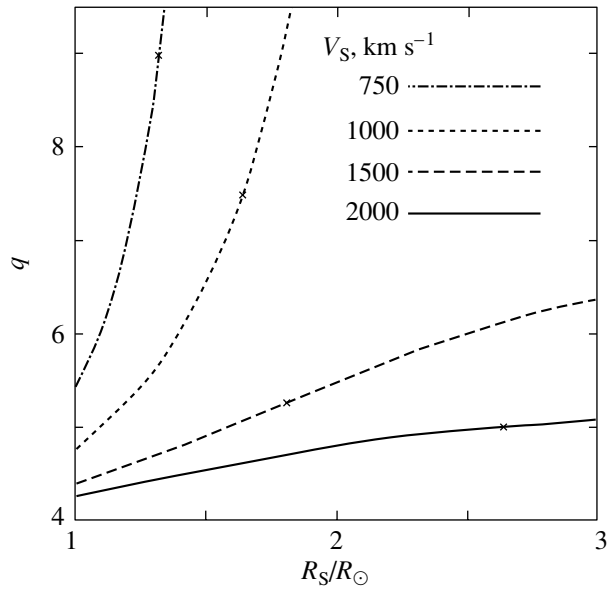

Fig. 3. The spectral index of the SCR spectrum at the shock front versus its radius R_{S} for various shock velocities V_{S}.

slowly changes with time (Berezhko *et al.* 1996, 1998). As applied to the case under consideration, formula (26) is useful for a qualitative analysis and rough estimation of p_{max}, because it disregards the significant changes in $q(R_{\text{S}})$.

Formula (26) establishes a relationship between the maximum momentum and the parameters of the problem, which in the nonrelativistic case is

$$p_{\text{max}} \propto \left(E_0 u_1/R_{\text{S}}^{\delta - 2\beta - 2}\right)^{1/(2 - \beta)}. \qquad (27)$$

Thus, in particular, we see that the maximum momentum decreases with increasing R_{S}, which results in the so-called runaway effect (Berezhko *et al.* 1996). Its essence is that at each time of shock evolution $t > 0$, the particles with momenta $p > p_{\text{max}}(t)$ that were produced at preceding stages when p_{max} was larger than the current value of $p_{\text{max}}(t)$ are increasingly accumulated in the upstream region. The propagation of these particles is weakly affected by the shock; the volume occupied by these particles due to their diffusion increases more rapidly than the shock size. For this reason, they are called escaping particles.

Escaping particles show up in the spectrum $J(\varepsilon)$ as an increasingly hard part near the maximum energy of $\varepsilon \sim 100$ MeV (see Fig. 2a). The formation of this high-energy bump is also facilitated by the fact that the purely power-law portion of the spectrum $J(\varepsilon) \propto \varepsilon^{-\gamma}$ at low energies, which consists of freshly accelerated particles, becomes increasingly soft with increasing R_{S} because of the decrease in the

Alfvén Mach number (see Fig. 3). For these reasons (the steepening of the spectrum of freshly accelerated particles and the decrease in their maximum energy starting from a certain time), the contribution of subsequent evolutionary stages to the resulting spectrum of the produced SCRs $N(\varepsilon)$ becomes negligible starting from some time t_f.

As we see from Fig. 2, the increase in the maximum energy ε_{max} and the change in the resulting SCR spectrum $N(\varepsilon)$ greatly slow down even when $R_S = 1.2R_\odot$ and virtually cease when the shock size reaches $R_S = 1.63R_\odot$ at $t_f = 434$ s.

In contrast to the current SCR spectrum at the shock front $J(\varepsilon)$, the resulting spectrum $N(\varepsilon)$ has a simpler shape that can be roughly represented as

$$N(\varepsilon) \propto \varepsilon^{-\gamma} \exp\left[-\left(\varepsilon/\varepsilon_{max}\right)^{\alpha}\right]. \quad (28)$$

The parameter α is difficult to predict analytically, because the index q changes significantly in the acceleration region. Analysis of the computed spectra $N(\varepsilon)$ shown in Fig. 2b indicates that $\alpha \simeq 2.5 - \beta$ and that the maximum energy is $\varepsilon_{max} = p_{max}^2/2m = 116, 75$, and 16 MeV for $\beta = 0.5, 1$, and 1.5, respectively. Formula (26) gives a reasonable estimate of the maximum energy: thus, for $\beta = 0.5$, it gives $\varepsilon_{max} \simeq 100$ MeV (in this case, $R_S = 1.2R_\odot$ should be taken).

The spectral index is $\gamma \simeq 2$ in all cases and the exponential part of the spectrum is smoother at large β, in accordance with expression (28). For this reason, the total energy content of the accelerated SCRs is larger for smaller β: $E_c = 2.1 \times 10^{30}, 1.8 \times 10^{30}$, and 1.1×10^{30} erg for $\beta = 0.5, 1$, and 1.5, respectively.

For technical reasons, the solution of the problem is currently very difficult to extend to much later evolutionary stages $t \gg t_f$ until the bulk of the SCRs produced in the solar corona reaches the Earth's orbit to directly compare calculations with SCR flux measurements, which are most often made near the Earth's orbit. The existence of the escaping effect allows this difficulty to be overcome. It may be assumed that at t_f all of the SCRs produced in the solar corona propagate purely diffusively into outer space without being appreciably affected by the shock.

Figure 4a shows the calculations of the resulting SCR spectrum $N(\varepsilon)$ for the same shock velocity $V_S = 1000$ km s^{-1} but in the case where the Alfvén turbulence decreases with radial distances more slowly, $\delta = 6$. Comparison with Fig. 2b shows that decreasing the parameter δ causes the maximum energy ε_{max} to increase, because a decrease in δ causes the turbulence level to increase at $r > R_\odot$.

In accordance with relation (27), the increase in the maximum energy ε_{max} is more significant at larger β: at $\delta = 6$, $\varepsilon_{max} = 172, 124$, and 38 MeV for $\beta = 0.5, 1$, and 1.5, respectively (see Fig. 4a).
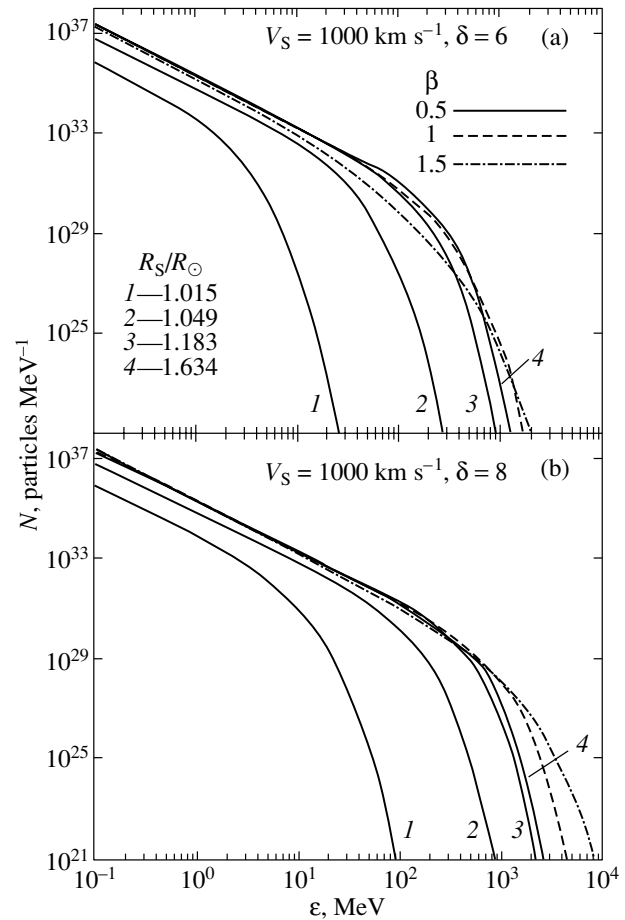

Fig. 4. Same as Fig. 2b for (a) $\delta = 6$ and (b) $E_0 = 3.7 \times 10^{-3}$ erg cm^{-3}.

Figure 4b shows the calculations of $N(\varepsilon)$, which differ from the case of Fig. 2b by a factor of 3.2 higher turbulence level. A comparison of Figs. 2b and 4b indicates that increasing the amplitude of the Alfvén turbulence spectrum E_0 causes the maximum energy ε_{max} to increase in the spectrum $N(\varepsilon)$, because the diffusion coefficient $\kappa_\parallel$ is inversely proportional to E_0. Therefore, from relation (25), one would expect the maximum momentum $p_{max} \propto E_0^{1/(2-\beta)}$ to increase with E_0; the larger the spectral index β, the larger the increase, as confirmed by our calculations: for the case shown in Fig. 4b, $\varepsilon_{max} = 460, 548$, and 494 MeV for $\beta = 0.5, 1$, and 1.5, respectively.

On the one hand, a change in the shock velocity V_S leads to the same effect as does a change in E_0, because $p_{max} \propto u_1^{1/(2-\beta)}$. On the other hand, the increase in velocity is accompanied by the formation of a harder power-law portion of the accelerated-particle spectrum at the same distances R_S because of the increase in the Alfvén Mach number: as we see from Fig. 3, the spectral index q significantly decreases
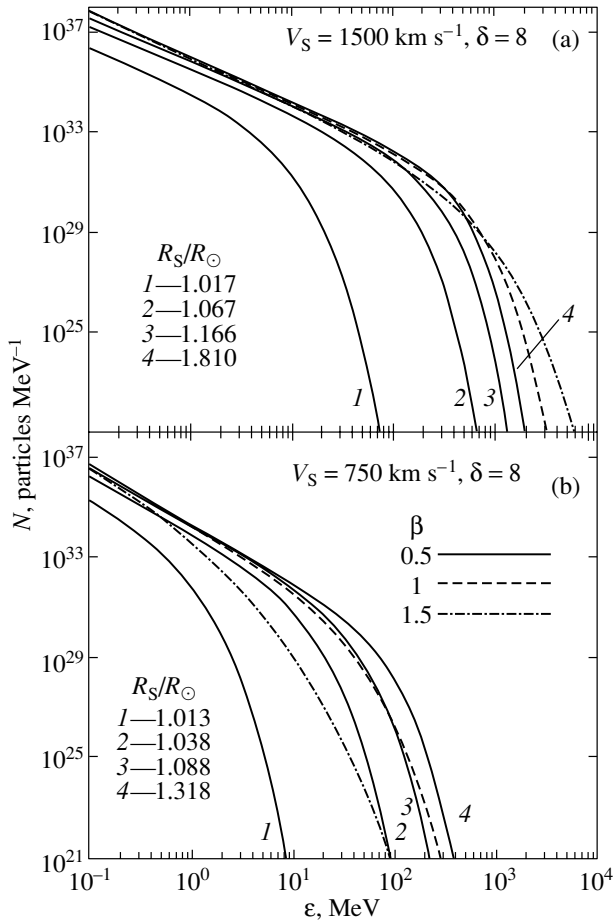
Fig. 5. Same as Fig. 2b for (a) $V_S = 1500$ km s^{-1} and $E_0 = 8 \times 10^{-4}$ erg cm^{-3} and (b) $V_S = 750$ km s^{-1} and $E_0 = 1.6 \times 10^{-3}$ erg cm^{-3}.

with increasing V_S. Therefore, one would expect the resulting SCR spectrum $N(\varepsilon)$ to be much more sensitive to V_S than to E_0.

This is confirmed by the calculations shown in Fig. 5a for $V_S = 1500$ km s^{-1} and $E_0 = 8 \times 10^{-4}$ erg cm^{-3}. Although the turbulence level is a factor of 1.5 lower than that in the case of Fig. 2b, the maximum SCR energy is much higher: $\varepsilon_{\max} = 227$, 227, and 136 MeV for $\beta = 0.5$, 1, and 1.5, respectively.

At the same time, increasing the shock velocity only by a quarter, as a comparison of Figs. 2b and 5b (where the computed resulting SCR spectrum $N(\varepsilon)$ is shown for $V_S = 750$ km s^{-1} and $E_0 = 1.6 \times 10^{-3}$ erg cm^{-3}) indicates, causes the maximum energy $\varepsilon_{\max}$ to significantly decrease: $\varepsilon_{\max} = 48$, 16, and 1.5 MeV for $\beta = 0.5$, 1, and 1.5, respectively.

The calculations presented above allow the back-reaction of the accelerated particles on the shock-

transition structure to be estimated. The main physical factor that determines the nonlinear shock modification is the accelerated-particle pressure

$$P_c = \frac{4\pi}{3} \int\limits_{p_{\mathrm{inj}}}^{\infty} p^3 v f \, dp. \qquad (29)$$

Using expression (22) for the distribution function at the shock front and taking into account the fact that, in our case, $q > 5$, $p_{\max} \gg p_{\mathrm{inj}}$, and $p_{\mathrm{inj}} \simeq 2mu_1$, we obtain

$$P_c = \frac{4\eta q}{3(q-5)} \rho_1 u_1^2. \qquad (30)$$

Substituting in the typical index $q = 6$ yields the ratio of the SCR pressure to the dynamic pressure of the shock $P_c/(\rho_1 u_1^2) = 8\eta$. Thus, we see that one might expect the shock to be significantly modified by the pressure of the shock-accelerated particles at injection rates much higher than $\eta = 10^{-3}$. As we show below, it follows from our comparison of calculations with experimental data that the required injection rates are $\eta \lesssim 10^{-5}$. This implies that $P_c \ll \rho_1 u_1^2$ and, hence, the shock modification by the accelerated-SCR pressure is negligible.

Another important aspect of the backreaction of the accelerated particles on the shock is the excitation of Alfvén waves in the upstream region $r > R_S$ by the SCR streaming instability. Based on the linear growth rate of the Alfvén waves, we can easily determine their energy density at the shock front (see, e.g., Lucek and Bell 2000):

$$E'_w = E_B \frac{u_1}{c_{a1}} \frac{P_c}{\rho_1 u_1^2}, \qquad (31)$$

where $E_B = B^2/8\pi$ is the energy density of the regular magnetic field. Taking into account the above estimate for the SCR pressure and the fact that most of the SCRs are produced at a stage when $M_{a1} \simeq 2$, we obtain

$$E'_w/E_B = 16\eta. \qquad (32)$$

The influence of the waves excited by SCR particles on the acceleration is determined by the relationship between their amplitude and the background turbulence amplitude. According to expressions (14) and (15), the energy content of the background turbulence at a shock front of radius R_S can be represented as

$$\frac{E_w}{E_B} = \frac{8\pi E_0}{\beta B_\odot^2} \left(\frac{R_S}{R_\odot} \right)^{-4}. \qquad (33)$$

Hence, given that $E_0/\beta \simeq 10^{-3}$ erg cm^{-3} and $B_\odot = 2.3$ G, the energy ratio of the excited and background

waves is

$$\frac{E'_w}{E_w} = 3 \times 10^3 \eta \left(\frac{R_S}{R_\odot}\right)^4. \qquad (34)$$

Thus, we see that the role of the turbulence excited by SCR particles depends on the injection rate η. The critical injection rate η_* that corresponds to the condition $E'_w/E_w = 1$ is given by the expression

$$\eta_* = 3 \times 10^{-4} \left(R_S/R_\odot\right)^{-4}. \qquad (35)$$

Since the maximum shock sizes lie within the range from $1.3R_\odot$ to $2.7R_\odot$ when the shock velocity V_S changes from 750 to 2000 km s^{-1}, we conclude that the critical injection rate η_* changes from 10^{-4} for $V_S = 750$ km s^{-1} to 6×10^{-6} for $V_S = 2000$ km s^{-1}.

As we show below, $\eta \leq 10^{-5}$ satisfy the experimental requirements. It thus follows that the excitation of Alfvén waves by accelerated particles can play a role only at extremely high shock velocities, $V_S > 1000$ km s^{-1}. Since the lowest-energy particles with $p \sim p_{\text{inj}}$ mainly contribute to the pressure P_c because of the steep SCR spectrum, one might expect that the excitation of long waves interacting with the highest-energy particles ($\varepsilon \sim \varepsilon_{\max}$) is negligible in all cases and cannot cause a significant increase in the maximum energy $\varepsilon_{\max}$.

A detailed conclusion about the role of the self-consistent generation of Alfvén waves during SCR acceleration in the solar corona can be drawn from quasi-linear calculations, which we are planning in the immediate future.

COMPARISON OF CALCULATIONS WITH EXPERIMENTAL DATA

Using the computed resulting spectrum $N_g(p)$ of the SCRs accelerated in the solar corona and the simplified procedure described in the section entitled "The SCR propagation", we can calculate the expected SCR fluxes $J(r, \varepsilon, t)$ far from the Sun and compare them with the measurements for specific events. The difficulties arising in such a comparison lie in the fact that several important parameters that significantly affect the SCR acceleration are not known well.

Thus, for example, the shock velocity, which was shown above to play a crucial role in the formation of the SCR spectrum, is known only in some cases (to be more precise, information on the CME velocity that must directly determine the shock velocity can be obtained from observations). It is also unclear to what extent the Alfvén turbulence level and spectral shape in the solar corona are subjected to variability from event to event. In addition, as was noted above, the injection rate η cannot currently be predicted with

the required accuracy. In our case of the linear theory, η affects only the amplitude of the SCR spectrum.

Despite these difficulties, it is of interest to compare the results of calculations with experimental data. The goal of this comparison is to determine whether the theory can reproduce the observed SCR fluxes and the shape of the SCR energy spectrum for the expected shock velocities of 500–2000 km s^{-1} and for Alfvén turbulence spectra that agree with the general energy constraints.

In our comparison with experiments, the parameters that determine the spectrum of the SCRs accelerated in the solar corona $N_f(\varepsilon)$ (shock velocity V_S, injection rate η, and turbulence level E_0) and the parameters β_e, δ_e, and E_{0e} that determine the Alfvén-wave spectrum at large distances $r > 10R_\odot$ and that directly affect the SCR propagation, were chosen so as to achieve the best agreement of the time variations in the SCR intensities $J(r, \varepsilon, t)$ at the point of observation $r = 1$ AU during the time interval Δt for the energies ε for which intensity J measurements are available. $\Delta t = t_2 - t_1$ corresponds to the time interval from the beginning of the event t_1 recorded by X-ray bursts in the $H\alpha$ line until the time $t_2 > t_1$ that was chosen in such a way that the interval Δt included the peak of the SCR intensity $J(r, \varepsilon, t)$ and, at the same time, was well ahead of the time of shock arrival at the point of observation in order to eliminate the influence of the shock-accelerated particles near the point of observation. Using the SCR intensity $J(r, \varepsilon, t)$ calculated in this way, we can determine the SCR spectrum at the point of observation

$$I(\varepsilon) = J_{\max}(\varepsilon), \qquad (36)$$

where $J_{\max}(\varepsilon) = J(r, \varepsilon, t_{\max})$ is the peak intensity of the SCRs with energy ε at the point of observation r and $t_{\max}(r, \varepsilon)$ is the time the peak value $J_{\max}$ is reached. As follows from relation (19), the intensity $I(\varepsilon)$ determined in this way is identical in shape to the spectrum $vN_f(\varepsilon)$ of the particles produced in the source (the corona): $I(\varepsilon) \propto vN_f(\varepsilon)$.

In Fig. 6, the SCR energy spectra for the September 29, 1989, event measured on IMP 8 and GOES and with neutron monitors (Lovell *et al.* 1998) are compared with the calculations.

The theoretical spectra were computed for a set of parameters in the acceleration region (in the corona)—$V_S = 2000$ km s^{-1}, $E_0 = 6 \times 10^{-4}$ erg cm^{-3}, $\delta = 8$, $\eta = 4.2 \times 10^{-6}$—and at large heliocentric distances $r > 10R_\odot$—$E_{0e} = 5 \times 10^{-14}$ erg cm^{-3}, $\beta_e = 1$, $\delta_e = 4$, $\Delta t = 0.8$ days. We see from the figure that the shape of the SCR spectrum at energies $\varepsilon < 300$ MeV depends only slightly on the spectral index β. At higher energies $\varepsilon > 300$ MeV, the SCR spectrum is much harder
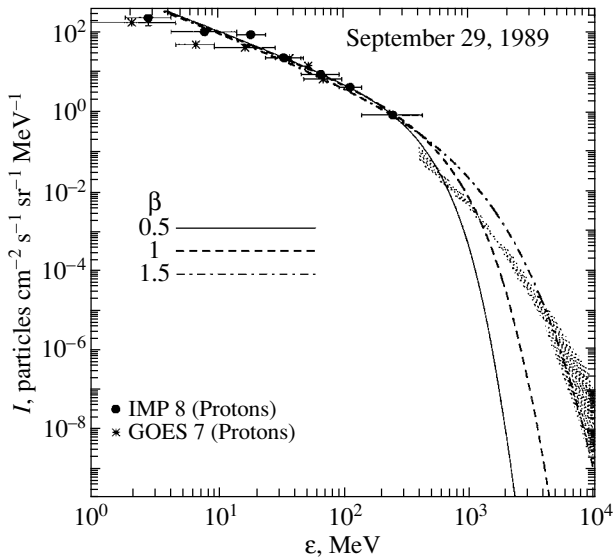
Fig. 6. The differential SCR intensity near the Earth's orbit versus the kinetic energy for the September 29, 1989 event. The IMP 8, GOES 7, and neutron-monitor (shaded region) measurements were taken from Lovell *et al.* (1998). The calculations were performed for $V_S = 2000$ km s^{-1}; $E_0 = 6 \times 10^{-4}$ erg cm^{-3}; $\delta = 8$; $\beta = 0.5$; 1, and 1.5, which correspond to $\eta = 4.2 \times 10^{-6}$; as well as $E_{0e} = 5 \times 10^{-14}$ erg cm^{-3}; $\beta_e = 1$; $\delta_e = 4$; and $\Delta t = 0.8$ days.
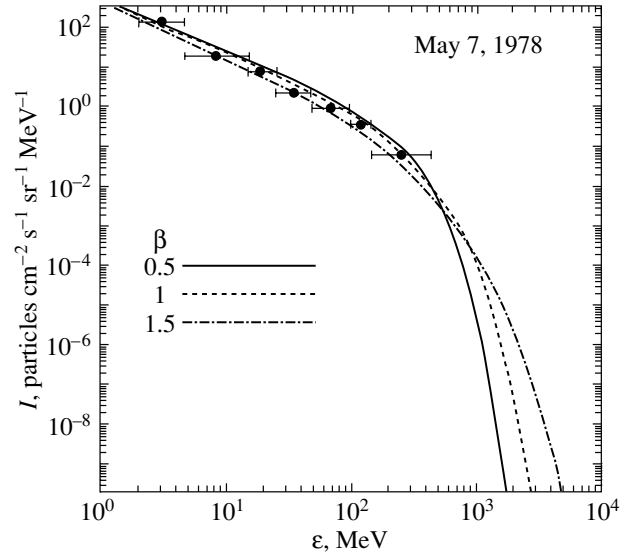

Fig. 7. The differential SCR intensity near the Earth's orbit versus the kinetic energy for the May 7, 1978, event. The IMP-8 measurements and the calculations performed for $V_S = 1500$ km s^{-1}; $E_0 = 8 \times 10^{-4}$ erg cm^{-3}; $\delta = 8$; $\beta = 0.5$, 1, and 1.5, which correspond to $\eta = 9 \times 10^{-7}$; as well as $E_{0e} = 5 \times 10^{-14}$ erg cm^{-3}; $\beta_e = 0.5$; $\delta_e = 2.5$; and $\Delta t = 1$ day, are shown.

for large β. It may be concluded with an allowance, made for the significant uncertainties in ground-based measurements, that calculations with $\beta = 1$ and 1.5 satisfactorily agree with the experimental data for quite a reasonable set of the remaining physical parameters. A similar situation takes place for the May 7, 1978, event, for which the SCR spectrum is shown in Fig. 7. The calculations were performed for $V_S = 1500$ km s^{-1}, $E_0 = 8 \times 10^{-4}$ erg cm^{-3}, $\delta = 8$, $\eta = 9 \times 10^{-7}$, $E_{0e} = 5 \times 10^{-14}$ erg cm^{-3}, $\beta_e = 0.5$, $\delta_e = 2.5$, and $\Delta t = 1$ day. As we see, in the energy range $\varepsilon < 300$ MeV in which IMP-8 measurements are available (http://hurlbut.jhuapl.edu/IMP/data/imp8/cpme/cpme_1h/protons/flux/), the computed spectra depend weakly on β and agree well with the experimental data.

Since the shock velocity required for agreement with the experimental data in this case is lower than that in the previous case, the SCR spectrum at relativistic energies $\varepsilon \gtrsim 300$ MeV is much steeper. Analysis of the responses of ground-based facilities indicates (Miroshnichenko 2001) that the total SCR flux at $\varepsilon > 435$ MeV is $I \simeq 1$ cm^{-2} s^{-1} sr^{-1}, a value that is nearly equal to the calculated value, which is almost independent of β.

A strong dependence of the observed SCR fluxes on the CME velocity is a distinctive feature of gradual

events. This dependence is illustrated by Fig. 8, in which the measured peak SCR fluxes $J_{\max}(\varepsilon)$ at energies $\varepsilon = 2$ MeV (Fig. 8a) and $\varepsilon = 20$ MeV (Fig. 8b) are plotted against the CME velocity V_p. This figure also shows the results of our calculations at $V_S \geq 400$ km s^{-1} for $\eta = 10^{-5}$ and $\beta = 0.5, 1, 1.5$.

Since the relationship between the CME velocity V_p and the CME-driven shock velocity V_S is not known from observations, we used the relation

$$V_S = V_p \sigma / (\sigma - 1), \qquad (37)$$

which is valid for a piston moving at a constant velocity V_p.

If we assume that the spectral indices of the Alfvén turbulence are different for different events, $\beta = 0.5-1.5$, then we may conclude that the calculations satisfactorily agree with the observations. If, however, we assume that the shape of the Alfvén turbulence spectrum in the solar corona is stable, then the spread in peak SCR fluxes at a fixed shock velocity can be attributed to the variability of the injection rate η. The latter possibility seems preferred, because the values of η lie in the range $10^{-6}-10^{-5}$ when analyzing individual events, which can be explained by the influence of the coronal magnetic-field structure.

As regards the shock velocities $V_S < 750$ km s^{-1}, the computed SCR fluxes rapidly decrease with decreasing V_S and lie well below the measured fluxes shown in Fig. 8 for $V_S \leq 500$ km s^{-1}. Such
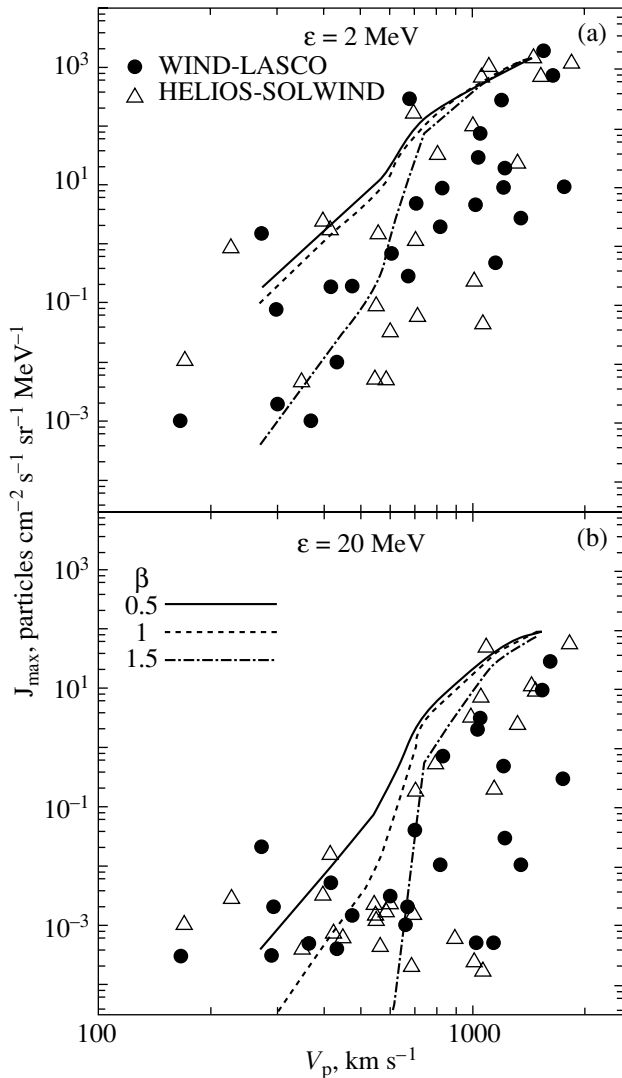
Fig. 8. The peak intensity of the SCRs with energies of (a) $\varepsilon = 2$ MeV and (b) $\varepsilon = 20$ MeV versus the CME velocity. The WIND and HELIOS (Kahler *et al.* 1984) measurements are shown. The lines represent the calculations of the SCR intensities expected at the Earth's orbit for $V_S = 400$, 750, 1000, 1500, and 2000 km s^{-1} at $E_0 = 1.2 \times 10^{-3}$ erg cm^{-3}; $\delta = 8$; $\beta = 0.5$, 1, and 1.5, which correspond to $\eta = 10^{-5}$; as well as $E_{0e} = 5 \times 10^{-14}$ erg cm^{-3}; $\beta_e = 0.5$; $\delta_e = 4$; and $\Delta t = 1.5$ days.

a strong dependence of the SCR production on V_S can be explained by the relatively high Alfvén velocity in the solar corona: as we see from Fig. 1, $c_a = 300$ km s^{-1} even at the base of the corona. As the Alfvén Mach number $M_a = (V_S - c_a - w)/c_a$ decreases, the accelerated-particle spectrum becomes increasingly steep (see Fig. 3), so the production of particles with energies $\varepsilon > 10$ MeV becomes inefficient at $V_S < 500$ km s^{-1}. Thus, the presence of measured SCR fluxes (although their number

is small) in Fig. 8 at $V_p < 500$ km s^{-1} cannot be described by our theory.

This inconsistency can be resolved by assuming that the Alfvén velocity c_a undergoes variations on the solar surface near the base of the corona. The experimental data in Fig. 8 at $V_p < 500$ km s^{-1} can be explained by assuming that there are local regions where c_a is a factor of 2 or 3 lower (e.g., because of the lower magnetic-field strength).

The calculations for $V_p = 270$ km s^{-1} (the shock velocity is $V_S = 400$ km s^{-1}) (Fig. 8) were performed for $B_\odot = 1.15$ G with the same values of the remaining parameters as those in the previous case. In this case, the Alfvén velocity is a factor of 2 lower than that in Fig. 1. We see from Fig. 8 that the assumption of this kind can account for the SCR fluxes observed at low CME velocities V_p.

On the other hand, according to Reames *et al.* (1997), only coronal mass ejections whose velocities exceed 500 km s^{-1} generate SCRs; those of them whose velocities are higher than 750 km s^{-1} are always accompanied by SCR events. This behavior closely matches the predictions of the theory presented above.

CONCLUSIONS

Our analysis of the SCR generation in the solar corona is based on the linear theory of diffusive shock acceleration, which disregards the backreaction of the accelerated-particle pressure on the shock-transition structure and the generation of Alfvén waves by the SCR streaming instability in the upstream region. This approximation is justified, on the one hand, by the small Alfvén Mach numbers ($M_a < 4$) in the coronal region where the bulk of the SCRs are produced and, on the other hand, by the low injection rate ($\eta \lesssim 10^{-5}$) required to produce the observed SCR fluxes.

The steep spectrum of the accelerated SCRs is the reason why the SCR pressure at all reasonable injection rates $\eta < 10^{-3}$ is negligible in comparison with the dynamic pressure in the shock. Accordingly, the shock modification by the accelerated-particle pressure is negligible.

As our analysis showed, the generation of Alfvén waves by accelerated particles can be significant either at an injection rate much larger than $\eta = 10^{-5}$ or at extremely high shock velocities $V_S \simeq 2000$ km s^{-1}. However, even in the latter case, intense generation of Alfvén waves is expected only at short wavelengths. At the same time, the excitation of the waves that interact with particles of extremely high energies $\varepsilon \sim \varepsilon_{max}$ is negligible in all cases.

Our calculations for realistic parameters of the solar corona and for the observed range of shock velocities indicate that an accelerated-particle spectrum that well agrees with the experimental requirements is formed when the shock propagates in the region $r < 2-3R_\odot$. In this case, the spectral shape of the produced SCRs is mainly determined by two physical parameters: the shock velocity V_S and the spectral index of the Alfvén turbulence β. Comparisons of the computed SCR spectra with those observed near the Earth's orbit for several individual events that show $\beta = 1-1.5$, which is consistent with available theoretical models, satisfy the experimental requirements. The shock velocity strongly affects both the total number of shock-produced SCRs and their maximum energy, in good agreement with the experiment.

Thus, our calculations lead us to conclude that the main observed features of the SCRs in the class of gradual events can be satisfactorily explained in terms of the diffusive shock acceleration theory.

ACKNOWLEDGMENTS

REFERENCES

1. V. E. Andreev, A. I. Efimov, L. N. Samoznaev, *et al.*, Solar Phys. **176**, 387 (1997).

2. W. I. Axford, E. Leer, and G. Skadron, *Proc. of the 15th Int. Conf. on Cosmic Rays, Plovdiv* **11**, 132 (1977).

3. E. G. Berezhko, Astropart. Phys. **5**, 367 (1996).

4. E. G. Berezhko, V. K. Elshin, G. F. Krymskiĭ, and S. I. Petukhov, *Generation of Cosmic Rays in Shock Waves* (Nauka, Novosibirsk, 1988).

5. E. G. Berezhko, V. K. Elshin, and L. T. Ksenofontov, Zh. Éksp. Teor. Fiz. **109**, 3 (1996) [JETP **82**, 1 (1996)].

6. E. G. Berezhko and G. F. Krymskiĭ, Usp. Fiz. Nauk **154**, 49 (1988) [Sov. Phys. Usp. **31**, 27 (1988)].

7. E. G. Berezhko, S. I. Petukhov, and S. N. Taneev, Pis'ma Astron. Zh. **24**, 151 (1998) [Astron. Lett. **24**, 122 (1998)].

8. E. G. Berezhko, S. I. Petukhov, and S. N. Taneev, Izv. Ross. Akad. Nauk, Ser. Fiz. **65**, 339 (2001a).

9. E. G. Berezhko, S. I. Petukhov, and S. N. Taneev, *Proc. 27-th Int. Conf. on Cosmic Rays, Hamburg* **8**, 3215 (2001b).

10. J. T. Burkepile and O. C. St. Cyr, *A Revised and Expanded Catalogue of Mass Ejections Observed by the Solar Maximum Mission Coronagraph* (High Altitude Observatory National Center for Atmospheric Research, Boulder Colorado, 1993).

11. D. C. Ellison, M. G. Baring, and F. C. Jones, Astrophys. J. **453**, 873 (1995).

12. D. C. Ellison and R. Ramaty, Astrophys. J. **298**, 400 (1985).

13. A. Hundhousen, *Coronal Expansion and Solar Wind* (Springer, Heidelberg, 1972; Mir, Moscow, 1976).

14. S. W. Kahler, J. T. Burkepile, and D. V. Reames, *Proc. 26th Int. Conf. on Cosmic Rays, Salt Lake City* **6**, 248 (1999).

15. S. W. Kahler, E. Hildner, and M. A. I. van Hollebeke, Solar Phys. **57**, 429 (1978).

16. S. W. Kahler, N. R. Sheeley, R. A. Howard, *et al.*, J. Geophys. Res. **89**, 9683 (1984).

17. J. Kota, E. Merenyi, J. R. Jokipii, *et al.*, Astrophys. J. **254**, 398 (1982).

18. S. M. Krimigis, J. Geophys. Res. **70**, 2943 (1965).

19. G. F. Krymskii, Dokl. Akad. Nauk SSSR **234**, 1306 (1977) [Sov. Phys. Dokl. **22**, 327 (1977)].

20. M. A. Lee, J. Geophys. Res. **87**, 5063 (1982).

21. M. A. Lee, J. Geophys. Res. **88**, 6109 (1983).

22. M. A. Lee and J. M. Ryan, Astrophys. J. **303**, 829 (1986).

23. J. L. Lovell, M. L. Duldig, and J. E. Humble, J. Geophys. Res. **103**, 23733 (1998).

24. S. G. Lucek and A. R. Bell, Mon. Not. R. Astron. Soc. **314**, 65 (2000).

25. M. A. Malkov and H. J. Völk, Astron. Astrophys. **300**, 605 (1995).

26. L. I. Miroshnichenko, *Solar Cosmic Rays* (Kluwer Acad., Dordrecht, 2001).

27. D. V. Reames, AIP Conf. Proc. **374**, 35 (1996).

28. D. V. Reames, Space Sci. Rev. **90**, 413 (1999).

29. D. V. Reames, AIP Conf. Proc. **516**, 289 (2000).

30. D. V. Reames, S. W. Kahler, and C. K. Ng, Astrophys. J. **491**, 414 (1997).

31. C. T. Russell, *Solar Wind*, Ed. by C. P. Sonett *et al.* (NASA SP-308, Washington, 1972), p. 365.

32. M. Scholer, K. J. Trattner, and H. Kucharek, Astrophys. J. **395**, 675 (1992).

33. E. C. Sittler and M. Guhathakurta, Astrophys. J. **523**, 812 (1999).

34. I. N. Toptygin, *Cosmic Rays in Interplanetary Magnetic Fields* (Nauka, Moscow, 1983).

35. K. J. Trattner and M. Scholer, Ann. Geophys. **9**, 774 (1993).

36. K. J. Trattner, E. Möbius, M. Scholer, *et al.*, J. Geophys. Res. **99**, 13 389 (1994).

37. C.-Y. Tu and E. Marsch, Space Sci. Rev. **73**, 1 (1995).

38. M. Velli and F. Pruneti, Plasma Phys. Control. Fusion **39**, B317 (1997).

39. You Qiu Hu, Shadia Rifai Habbal, and Xing Li, J. Geophys. Res. **104**, 24 819 (1999).

40. G. P. Zank, W. K. M. Rice, and C. C. Wu, J. Geophys. Res. **105**, 25 079 (2000).

Translated by V. Astakhov

Using Wavelet Approximation to Study the Influence of Jupiter on the Orbital Evolution of a Distant Satellite of Saturn

M. A. Vashkov'yak*

Keldysh Institute of Applied Mathematics, Russian Academy of Sciences,
Miusskaya pl. 4, Moscow, 125047 Russia
Received March 5, 2003

Abstract—We suggest a nonstandard methodology for studying the influence of Jupiter on the secular orbital evolution of a distant satellite of Saturn. This influence is tangible only in short time spans near the times of the smallest separation between Jupiter and Saturn, i.e., when the heliocentric longitudes of the two planets coincide. These times are spaced about 20 years apart. To describe the jumplike behavior of perturbations, we suggest approximating the principal part of the perturbing function averaged over the satellite's motion by a two-parameter exponential wavelet-type (burst) function. The subsequent averaging (smoothing) of the perturbing function allows us to eliminate the 20-year-period terms and obtain an approximate analytical solution in a special case of the problem. The results are illustrated by plots of the variations in the averaged perturbing function and the orbital eccentricity of Saturn's outer satellite S/2000 S1, which is most strongly perturbed by Jupiter. © *2003 MAIK "Nauka/Interperiodica"*.

Key words: *celestial mechanics, orbital evolution, and wavelet approximation.*

INTRODUCTION

Recently, satellite systems that are far removed both from the giant planets themselves and from the orbits of their inner satellites have been discovered. The main force that perturbs the elliptical motion of the new outer satellites is solar attraction, which exceeds the combined effect of the largest inner satellites and the oblateness of the central planet by several orders of magnitude. The strongest planetary perturbations, in particular, of Saturn's satellites are attributable to the influence of only the most massive planet, Jupiter. In addition, these perturbations show up only near the times t_k ($k = 1, 2, \ldots$) of the smallest separation between Jupiter and Saturn. Such a "burstlike" behavior of the perturbations is difficult to describe by means of the commonly used Fourier series, because a large number of harmonics is required to approximate the perturbing function of Jupiter.

In this paper, we suggest a nonstandard procedure that uses a special two-parameter burst-type function and the basics of wavelet analysis (Chui 2001) to study the main perturbations from Jupiter. The parameters of this function are chosen to best approximate the actual perturbing function of the evolution problem. In particular, the times t_k that correspond to

the strongest perturbations from Jupiter for sequential values of $k = 1, 2, \ldots$ are spaced

$$T = \frac{T_1 T_2}{T_1 - T_2} \approx 20 \text{ years}$$

apart (here T_1 and T_2 are the orbital periods of Saturn and Jupiter, respectively). At these times, the heliocentric longitudes of Saturn, λ_1, and Jupiter, λ_2, satisfy the equality

$$\lambda_{2k} = \lambda_{1k} + 2k\pi. \qquad (1)$$

Below, to eliminate the terms with the 20-year period from the perturbing function W, we smooth or, in a sense, average the bursts of the perturbing function by a method similar to that used in wavelet analysis. The resulting averaged perturbing function $\overline{W}$ becomes (along with the semimajor axis of the satellite orbit) the first integral of the system of five evolution equations. When applied to the farthest satellite of Saturn, S/2000 S1, this method yielded an approximate analytical solution of the equation for the eccentricity variation because of the peculiar features of its orbit (its inclination is close to 180°). Compared to the eccentricity, the remaining orbital elements are only slightly perturbed by Jupiter.

FORMULATION OF THE PROBLEM

Let us consider the motion of a distant satellite of Saturn under the attraction of the Sun and Jupiter.

---

*E-mail: vashkov@keldysh.ru

We assume that Jupiter and Saturn move heliocentrically in circular orbits that lie in the plane of the ecliptic. The Saturnocentric motion of the satellite is elliptical and three-dimensional. Denote the heliocentric radius vectors of Saturn and Jupiter by $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively; the Saturnocentric radius vectors of the satellite and Jupiter by $\mathbf{r}$ and $\boldsymbol{\rho}$, respectively; and the masses of the Sun, Saturn, and Jupiter by m_j ($j = 0, 1,$ and 2, respectively). The mass of the satellite is assumed to be infinitesimal. Since the values of $r/r_1 \approx 0.016$ and $r/\rho \le 0.035$ are relatively small for the farthest discovered satellite of Saturn, S/2000 S1, we can retain only the principal terms in the perturbing function R of the attraction of the Sun and Jupiter, i.e., restrict our analysis to the Hill approximation. In this case,

$$R = R_1 + R_2, \qquad (2)$$

$$R_1 = \frac{fm_0}{2r_1^3}\left[\frac{3}{r_1^2}\left(\mathbf{r}, -\mathbf{r}_1\right)^2 - r^2\right], \qquad (3)$$

$$R_2 = \frac{fm_2}{2\rho^3}\left[\frac{3}{\rho^2}\left(\mathbf{r}, \rho\right)^2 - r^2\right], \qquad (4)$$

where $(\mathbf{x}, \mathbf{y})$ denotes the scalar product of the vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\mathbf{r} = \begin{pmatrix} r^{(1)} \\ r^{(2)} \\ r^{(3)} \end{pmatrix},$$

$$\mathbf{r_1} = a_1 \begin{pmatrix} \cos\lambda_1 \\ \sin\lambda_1 \\ 0 \end{pmatrix}, \quad \mathbf{r_2} = a_2 \begin{pmatrix} \cos\lambda_2 \\ \sin\lambda_2 \\ 0 \end{pmatrix},$$

$$\rho = \mathbf{r_2} - \mathbf{r_1} = \begin{pmatrix} a_2\cos\lambda_2 - a_1\cos\lambda_1 \\ a_2\sin\lambda_2 - a_1\sin\lambda_1 \\ 0 \end{pmatrix} \qquad (5)$$

$$= \rho \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix},$$

$$\rho^2 = a_1^2 + a_2^2 - 2a_1a_2\cos(\lambda_2 - \lambda_1). \qquad (6)$$

Here, f is the gravitational constant; a_1 and a_2 are the radii of the heliocentric orbits of Saturn and Jupiter, respectively; and φ is the Saturnocentric longitude of Jupiter. Our objective is to analyze the secular orbital evolution of the satellite and, primarily, the evolution of the eccentricity as the orbital element that is

most strongly perturbed by Jupiter. To facilitate the allowance for the influence of rare bursts of Jupiter's attraction on the orbital evolution, we perform a number of transformations of the perturbing function.

PRELIMINARY AVERAGING TRANSFORMATIONS

Let us eliminate the short-period terms related to the orbital motion of the satellite (the orbital period of S/2000 S1 is about 3.64 years) from the perturbing function. This elimination is achieved by averaging the function R over the mean longitude of the satellite.

To eliminate the longer-period terms requires averaging over the motion of the Sun (Saturn) and Jupiter. The former is relatively easy to perform in the absence of Jupiter ($m_2 = 0$) and yields the standard integrable doubly averaged Hill problem (Lidov 1961; Kozai 1962). This averaging procedure eliminates the terms with the orbital period of Saturn $T_1 \approx$ 29.5 years from the perturbing function.

As a result of the above transformations, we obtain the following preaveraged perturbing function:

$$W = \frac{1}{4\pi^2}\int\limits_0^{2\pi}\int\limits_0^{2\pi} R_1\,d\lambda d\lambda_1 + \frac{1}{2\pi}\int\limits_0^{2\pi} R_2\,d\lambda \qquad (7)$$

$$= \frac{3fm_0a^2}{16a_1^3}\Big\{(1+\nu)[2(e^2 - \sin^2 i)$$

$$+ e^2\sin^2 i(5\cos 2\omega - 3)] + \nu[2\sin^2 i$$

$$+ 10e^2\cos 2\omega + e^2\sin^2 i(3 - 5\cos 2\omega)]$$

$$\times \cos 2(\varphi - \Omega) + 10\nu e^2\cos i\sin 2\omega\sin 2(\varphi - \Omega)\Big\},$$

where a, e, i, ω, and Ω are the standard designations for the Keplerian orbital elements of the satellite; $\nu = \nu(t) = \dfrac{m_2}{m_0}\left[\dfrac{a_1}{\rho(t)}\right]^3$, $\varphi = \varphi(t)$, and $\rho = \rho(t)$ can be determined from formulas (5) and (6); and longitudes λ_1 and λ_2 are linear functions of time:

$$\lambda_j = \lambda_{j0} + \frac{2\pi}{T_j}(t - t_0), \quad (j = 1, 2). \qquad (8)$$

For $m_2 = \nu = 0$, formula (7) describes the perturbing function of the doubly averaged Hill problem. When $m_2 \ne 0$, the system of five evolution equations with the perturbing function W takes the form

$$\frac{da}{d\tau} = 0, \qquad (9)$$

$$\frac{de}{d\tau} = 10e\sqrt{1 - e^2}\left[(1 + \nu)\sin^2 i\sin 2\omega\right.$$

$$+ C\left(2 - \sin^2 i\right)\sin 2\omega - 2S\cos i\cos 2\omega\right],$$

$$\frac{di}{d\tau} = -\frac{2\sin i}{\sqrt{1 - e^2}}\left\{5\left(1 + \nu - C\right)e^2\cos i\sin 2\omega\right.$$

$$+ S \left[2 + e^2 \left(3 + 5 \cos 2\omega \right) \right] \},$$

$$\frac{d\omega}{d\tau} = \frac{2}{\sqrt{1 - e^2}} \Big\{ (1 + \nu) \left[4 + e^2 - 5 \sin^2 i \right.$$

$$\left. + 5 \left(\sin^2 i - e^2 \right) \cos 2\omega \right]$$

$$+ C \left[5 \left(2 - e^2 - \sin^2 i \right) \cos 2\omega \right.$$

$$\left. + 5 \sin^2 i - 2 - 3e^2 \right]$$

$$+ 5S(2 - e^2) \cos i \sin 2\omega \Big\},$$

$$\frac{d\Omega}{d\tau} = -\frac{2}{\sqrt{1 - e^2}}$$

$$\times \Big\{ (1 + \nu - C) \left[2 + e^2 \left(3 - 5 \cos 2\omega \right) \right] \cos i$$

$$+ 5Se^2 \sin 2\omega \Big\},$$

$$C = \nu \cos 2(\varphi - \Omega), \quad S = \nu \sin 2(\varphi - \Omega), \quad (10)$$

$$\tau = \frac{3m_0 a^3}{16 m_1 a_1^3} n \left(t - t_0 \right);$$

$n = \sqrt{\dfrac{f m_1}{a^3}}$ is the mean motion of the satellite; and the normalized time τ acts as an independent variable.

Since the function W depends on all of the orbital elements and explicitly on time and system (9) has only one first integral $a = \text{const}$, we solve it mainly numerically.

We numerically integrate system (9) with the initial data that correspond to the orbit of Saturn's satellite S/2000 S1 (Marsden 2001):

$$\text{JD}(t_0) = 2452200.5, \quad a = 0.1557901 \text{ AU},$$

$$e = 0.374658,$$

$$i = 172°.74556, \quad \omega = 42°.43394, \quad \Omega = 207°.06661.$$

The initial longitudes of Saturn and Jupiter are $\lambda_{10} \approx 72°$ and $\lambda_{20} \approx 89°$, respectively. We apply the fourth-order Runge−Kutta method, which is successfully used for the evolution problems that do not require calculating the position of the satellite in its orbit. The integration step automatically changed from 0.1 to 1 year.

The thin solid line in Fig. 1 indicates the time dependence of the function W normalized to $3f m_0 a^2 / 16/a_1^3$ and calculated using Eqs. (9) for a 250-year-long interval. For comparison, the dashed line marks the constant value of W_0 normalized in the same way, which is the first integral of the evolution system for $m_2 = 0$. Clearly, the perturbations from Jupiter are strongest near the times t_k, when condition (1) is satisfied. In this case,

$$\rho = \rho_{\min} = a_1 - a_2, \quad (11)$$

$$\varphi = \lambda_{2k} - \pi = \lambda_{1k} + 2k\pi - \pi,$$

$$\lambda_{1k} = \lambda_{2k} - 2k\pi$$

$$= (\lambda_{10} n_2 - \lambda_{20} n_1 + 2k\pi n_1)/(n_2 - n_1),$$

and the times t_k themselves are defined by the formula

$$t_k = t_0 + (\lambda_{10} - \lambda_{20} + 2k\pi)/(n_2 - n_1), \quad (12)$$

$$k = 1, 2, \ldots,$$

where $n_1 = 2\pi/T_1$ and $n_2 = 2\pi/T_2$ are the mean motions of Saturn and Jupiter, respectively. With the adopted λ_{10} and λ_{20}, $t_1 \approx t_0 + 18$ years and the remaining t_k are spaced about 20 years apart.

Note that, over 250 years, the amplitude of the perturbation bursts from Jupiter at times t_k may change by severalfold, being at a minimum at $t_4 \approx 78$ years and $t_5 \approx 98$ years. Qualitatively, these minima can be explained by the peculiar features of the orbit of S/2000 S1 and its evolution. Since the inclination of this orbit is close to $180°$ ($\sin^2 i \approx 0.016$), we may assume, to a first approximation, that $\cos i \approx -1$ and disregard the terms $\sim \nu \sin^2 i$ in the expression for the perturbing function (7). We then obtain

$$W \approx \frac{3f m_0 a^2}{16 a_1^3} \Big\{ 2 \left(e^2 - \sin^2 i \right) \quad (13)$$

$$+ e^2 \sin^2 i \left(5 \cos 2\omega - 3 \right)$$

$$+ 2\nu e^2 \left[1 + 5 \cos 2(\varphi - \Omega + \omega) \right] \Big\}.$$

Here, φ varies together with the longitudes λ_1 and λ_2, while Ω and ω evolve under the effect of solar perturbations. The last term $\sim \nu$ describes the perturbations from Jupiter. Having constructed the function

$$f(t) = 1 + 5 \cos 2[\varphi(t) - \Omega(t) + \omega(t)]$$

on the given time interval or, more precisely, having calculated this function using Eqs. (9), we can verify that two of its zeros are reached at times close to t_4 and t_5. This implies that the amplitudes of the corresponding bursts $W(t_4)$ and $W(t_5)$ relative to W_0 are close to zero (see Fig. 1).

Because of the burstlike variation in the perturbing function W, we can hardly hope to successfully approximate it by a manageable number of Fourier harmonics. Such an approximation can be obtained by using two-parameter burst functions similar to those used in wavelet analysis.

WAVELET APPROXIMATION OF THE PERTURBING FUNCTION

The simplest and roughest approximation of the function W could be realized by assuming the pulsed pattern of the perturbations that act only at times t_k. In this case, ρ, φ, and λ_{1k} should be formally substituted with the quantities defined by formulas (11), and
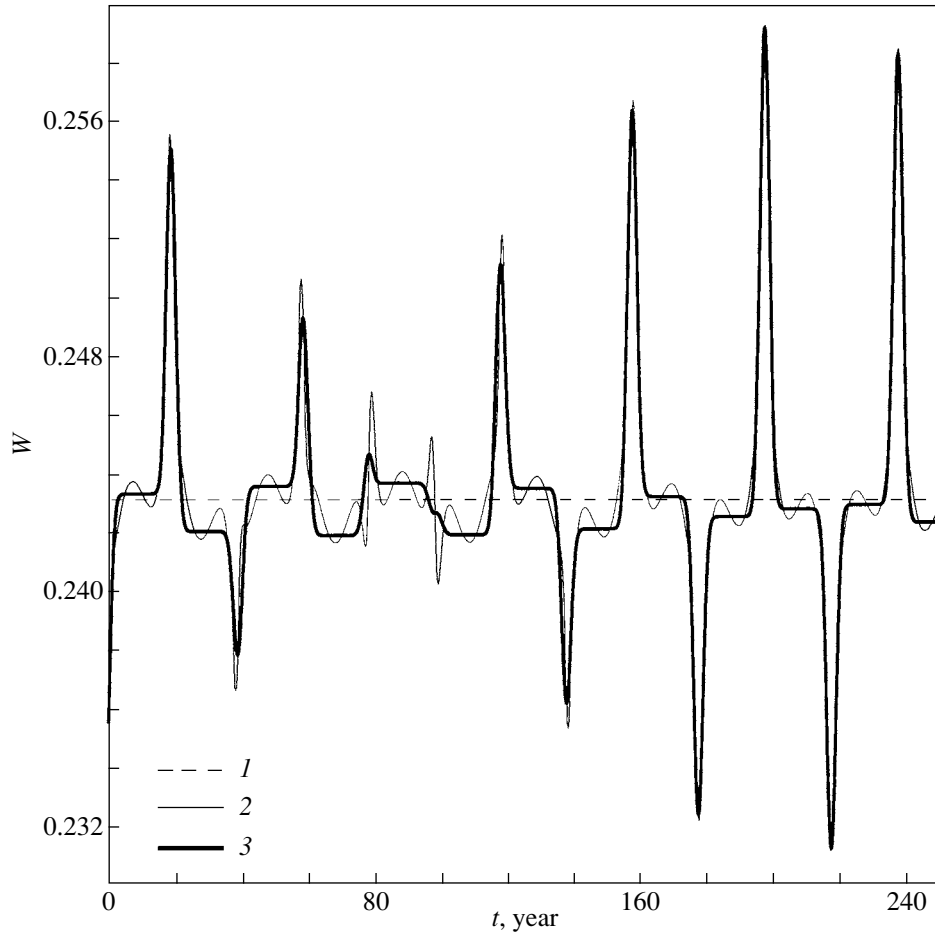
Fig. 1. Time dependences of the averaged perturbing function: (1) W_0 ($m_2 = 0$) = const; (2) W($m_2 \neq 0$); and (3) W_a.

the part of the perturbing function $W \sim \nu$ should be multiplied by the function

$$\delta(t - t_k) = \begin{cases} 1, & t = t_k \\ 0, & t \neq t_k, \end{cases} \quad (14)$$

where t_k is defined by formula (12). However, function (14) contains only one parameter and cannot accurately describe the perturbation bursts.

Actually, the perturbing effect of Jupiter is "smeared" in time relative to the times t_k. In addition, it also shows up (although much more weakly) between these times. It seems natural to describe such a burstlike process by a system of special two-parameter functions

$$\psi(t, s, \sigma) = \psi\left(\frac{t - s}{\sigma}\right). \quad (15)$$

We use the function

$$\psi = \exp\left[-\left(\frac{t - s}{\sigma}\right)^2\right], \quad (16)$$

which is suitable for approximating bell-shaped perturbations, and call the corresponding approximation a wavelet approximation. The parameters s and σ in function (16) are determined by the specific problem. The parameter s specifies the times of maximum action of the perturbations, and σ specifies the width of the bell-shaped function at height $\psi_\sigma = 1/e$ (here, $e = 2.718\ldots$). In our case, the parameters s and σ (in contrast to strict wavelet analysis) are, as it were, preassigned and constant: s are the times t_k and σ is the empirical value of two years chosen from the best approximation condition. As a result, the function W is described by the formula

$$\begin{aligned} W_a = &\frac{3fm_0a^2}{16a_1^3}\{(1 + \mu(t))[2(e^2 - \sin^2 i) \quad (17)\\ &+ e^2\sin^2 i(5\cos 2\omega - 3)] + \mu(t)[2\sin^2 i\\ &+ 10e^2\cos 2\omega + e^2\sin^2 i(3 - 5\cos 2\omega)]\\ &\times \cos 2(\lambda_{1k} - \Omega)\\ &+ 10\mu(t)e^2\cos i\sin 2\omega\sin 2(\lambda_{1k} - \Omega)\}, \end{aligned}$$

where

$$\mu(t) = \frac{m_2}{m_0}\left(\frac{a_1}{a_1 - a_2}\right)^3 \exp\left[-\left(\frac{t - t_k}{\sigma}\right)^2\right] \quad (18)$$

and λ_{1k} and t_k are defined by formulas (11) and (12), respectively. Formula (17) can be derived from Eq. (7) by formally substituting μ for ν and λ_{1k} for φ. In Fig. 1, the normalized function W_a is indicated by the heavy curve. This function satisfactorily approximates the maximum perturbation bursts and its deviations from the function W show up only between bursts ("wavelet segments") except for the neighborhoods of t_4 and t_5, where the perturbations from Jupiter are at a minimum. The difference in the levels of the horizontal segments of W_a stems from the fact that this function was calculated by using the differential equations that describe the variations in the orbital elements of the satellite. If their initial values are used to calculate the function W_a, then all horizontal segments will be at the level of W_0.

Below, we describe the variations in the orbital eccentricity of the satellite S/2000 S1. This orbital element, which characterizes the shape of the orbit, is most strongly affected by the perturbing influence of Jupiter, which clearly shows up against the background of secular solar perturbations. For better illustration, we restrict our analysis to a 90-year-long interval, which slightly exceeds the period of the long-period eccentricity variations due to solar perturbations (about 73 years). Figure 2 shows the eccentricity variations that were determined by numerically solving system (9). The dashed line corresponds to the solar perturbations alone ($m_2 = 0$). The solid lines represent the dependences $e(t)$ for $m_2 \neq 0$; the thin line corresponds to the equations with the perturbing function W, and the heavy line corresponds to the wavelet approximation of this function, W_a. The thin line can be barely distinguished from the heavy line only in some time intervals near $t \approx 17$, 37, and 57 years. The dependences indicated by the dots and filled triangles are explained in the next section.

THE AVERAGING TRANSFORMATION OF THE FUNCTION W_a

When the wavelet approximation is used, the part of the function W_a that is proportional to μ depends explicitly on time through $\exp\left[-\left(\frac{t - t_k}{\sigma}\right)^2\right]$; t_k, as well as λ_{1k}, is a piecewise constant function of the number $k = E\left(\frac{\lambda_2(t) - \lambda_1(t) + \pi}{2\pi}\right)$. Denoting the

difference between W_a and $W_0 = W(m_2 = 0) = $ const by $\Delta W(k)$, we obtain

$$\Delta W(k) = A_k(e, i, \omega, \lambda_{1k} - \Omega)\exp\left[-\left(\frac{t - t_k}{\sigma}\right)^2\right]. \quad (19)$$

Here,

$$A_k = \frac{3fm_2a^2}{16a_1^3}\left(\frac{a_1}{a_1 - a_2}\right)^3\{2(e^2 - \sin^2 i) \quad (20)$$
$$+ e^2\sin^2 i(5\cos 2\omega - 3) + [2\sin^2 i + 10e^2\cos 2\omega$$
$$+ e^2\sin^2 i(3 - 5\cos 2\omega)]\cos 2(\lambda_{1k} - \Omega)$$
$$+ 10e^2\cos i\sin 2\omega\sin 2(\lambda_{1k} - \Omega)\}.$$

The time-averaged value of the function $\Delta W(k)$ on the individual complete kth wavelet segment can be calculated using the formula

$$\overline{\Delta W(k)} = \frac{1}{2\beta\sigma}\int_{t_k - \beta\sigma}^{t_k + \beta\sigma} A_k \exp\left[-\left(\frac{t - t_k}{\sigma}\right)^2\right] dt, \quad (21)$$

where β is a positive number that is determined by the maximum separation between the opposite fronts of the burst or by its maximum width equal to $2\beta\sigma$.

In the example under consideration, this width is about eight years, or $\beta \approx 4$. The difference between $\Delta W(k)$ and the values corresponding to the horizontal parts of the function W is of the order of $\exp(-16) \approx 10^{-7}$ if $|t - t_k| \geq 4\sigma$ (or eight years); i.e., the methodological parameter β determines the burst boundaries beyond which the function ψ may be set equal to zero.

To approximately calculate the integral (21), we assume that the satellite orbital elements on which A_k depend are constant within the kth integration interval; i.e., A_k are time-independent. Therefore, we ignore the difference between the levels of the horizontal portions of the function $\Delta W(k)$. We then have

$$\overline{\Delta W(k)} = \frac{A_k}{\beta\sigma}\int_{t_k}^{t_k + \beta\sigma}\exp\left[-\left(\frac{t - t_k}{\sigma}\right)^2\right] dt \quad (22)$$

$$= \frac{A_k}{\beta}\int_0^{\beta}\exp\left[-\theta^2\right] d\theta = \frac{\sqrt{\pi}A_k}{2\beta}\Phi(\beta).$$

Here, $\theta = (t - t_k)/\sigma$; $\Phi(\beta)$ is a special function, the so-called error function. For $\beta = 4$, we have $\Phi(4) = 1 + O(10^{-7})$ (Abramowitz and Stegun 1964). Disregarding the deviation of $\Phi(4)$ from unity, we obtain

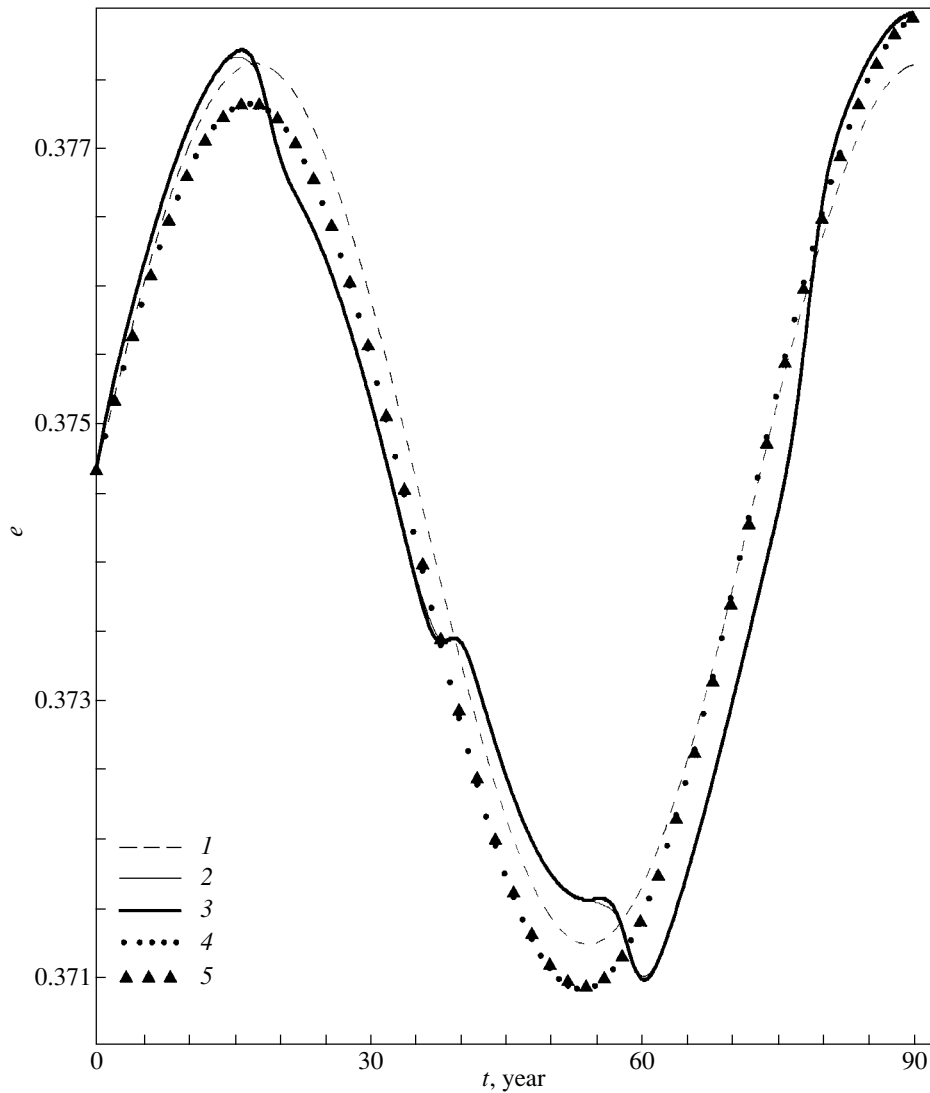$$\overline{\Delta W(k)} = \frac{\sqrt{\pi}A_k}{8}, \quad (23)$$

Fig. 2. The time dependence of the orbital eccentricity of the satellite S/2000 S1: (*1*) $m_2 = 0$; (*2*) $m_2 \neq 0$; (*3*) $m_2 \neq 0$, the wavelet approximation W_a; (*4*) $m_2 \neq 0$, the averaged function $\overline{W_a}$; and *5*, the approximate analytical solution.

where A_k is defined by formula (20).

The final averaging over all kth segments yields the function $\overline{\Delta W}$ as an arithmetic mean of the functions $\overline{\Delta W(k)}$:

$$\overline{\Delta W} = \frac{1}{K} \sum_{k=1}^{K} \overline{\Delta W(k)}, \qquad (24)$$

where K is the number of complete bursts (wavelet segments) in a specified time interval. For the 250- and 90-year intervals shown in Figs. 1 and 2, $K = 12$ and 4, respectively. When calculating the finite sums $\sum_{k=1}^{K} \frac{\sin}{\cos} (\lambda_{1k} - \Omega)$, we use formulas (11) and

the relations

$$\sum_{k=1}^{K} \frac{\sin}{\cos} (kx) = \frac{\sin\left(\dfrac{Kx}{2}\right)}{\sin\left(\dfrac{x}{2}\right)} \left\{ \frac{\sin}{\cos} \left(\frac{K+1}{2}x\right) \right\}. \qquad (25)$$

Assuming that $x = \dfrac{4\pi n_1}{n_2 - n_1}$, we obtain the following expression for the averaged perturbing function $\overline{W_a}$ (the averaging eliminates all of the terms with periods of 30, 20, and 3.64 years):

$$\overline{W_a} = W_0 + \overline{\Delta W} = \frac{3 f m_0 a^2}{16 a_1^3} \qquad (26)$$

$$\times \{(1 + \zeta)[2(e^2 - \sin^2 i) + e^2 \sin^2 i(5\cos 2\omega - 3)]$$

$$+ \eta[2\sin^2 i + 10e^2 \cos 2\omega$$
$$+ e^2 \sin^2 i(3 - 5\cos 2\omega)]\cos 2(\chi - \Omega)$$
$$+ 10\eta e^2 \cos i \sin 2\omega \sin 2(\chi - \Omega)\}.$$

In this expression, ζ, η, and χ are the constants defined by the formulas

$$\zeta = \frac{\sqrt{\pi}}{8}\frac{m_2}{m_0}\left(\frac{a_1}{a_1 - a_2}\right)^3, \quad \eta = \frac{\zeta}{K}\frac{\sin(Ky)}{\sin y}, \quad (27)$$

$$y = \frac{x}{2} = \frac{2\pi n_1}{n_2 - n_1},$$

$$\chi = \frac{1}{n_2 - n_1}\left[\lambda_{10}n_2 - \lambda_{20}n_1 + (K+1)\pi n_1\right].$$

The function $\overline{W_a}$ is the first integral of the evolution system that is derived from (9) by formally substituting ζ for ν, $\nu\cos 2(\varphi - \Omega)$ for $\eta\cos 2(\chi - \Omega)$, and $\eta\sin 2(\chi - \Omega)$ for $\nu\sin 2(\varphi - \Omega)$.

In this case, the constant of the integral is determined by K or the length of the time interval under consideration. As an example, we give the normalized values

$$\overline{W_a}(K = 4) = 0.24388,$$
$$\overline{W_a}(K = 12) = 0.24361, W_0 = 0.24313.$$

Note that $\overline{W_a}(K \to \infty) \to W_0$ particularly rapidly for alternating bursts.

The large dots in Fig. 2 indicate the variations in the orbital eccentricity of the satellite S/2000 S1 determined by numerically integrating the evolution system with the averaged perturbing function $\overline{W_a}$. This dependence smoothes the heavy line that corresponds to the unaveraged function W_a.

AN APPROXIMATE SOLUTION
FOR $\sin i \approx 0$

The two existing first integrals $a = \text{const}$ and $\overline{W_a} = \text{const}$ are not enough to find the general solution of the evolution system. However, the particular solution $\sin i = \sin i_0 = 0$ exists for $i_0 = 0, \pi$, while an approximate analytical solution can be obtained for the variation of the most perturbed orbital element (eccentricity) in the actual case of $\sin i \approx 0$.

Disregarding the terms $\sim \nu\sin^2 i$ in formula (26), we obtain, by analogy with (13),

$$\overline{W_a} \approx \frac{3fm_0a^2}{16a_1^3}\{2(e^2 - \sin^2 i) \quad (28)$$

$$+ e^2 \sin^2 i(5\cos 2\omega - 3)$$

$$+ 2e^2[\zeta + 5\eta\cos 2(\varphi - \Omega + \omega)]\}.$$

Let us assume that

$$e = \overline{e} + \delta e, \quad i = \overline{i} + \delta i, \quad (29)$$

$$\omega = \overline{\omega} + \delta\omega, \quad \Omega = \overline{\Omega} + \delta\Omega,$$

where $\overline{e}, \overline{i}, \overline{\omega}$, and $\overline{\Omega}$ are the functions of time that satisfy the system of evolution equations with $m_2 = 0$. In the averaged problem, the variation in eccentricity with time can be roughly described by the equation

$$\frac{de}{dt} = -\frac{\sqrt{1-e^2}}{ena^2}\frac{\partial\overline{W_a}}{\partial\omega} = \frac{15fm_0}{8a_1^3 n} \quad (30)$$

$$\times e\sqrt{1-e^2}\left[\sin^2 i \sin 2\omega + 2\eta\sin 2(\chi - \Omega + \omega)\right].$$

In this case,

$$\frac{d\overline{e}}{dt} = \frac{15fm_0}{8a_1^3 n}\overline{e}\sqrt{1-\overline{e}^2}\sin^2\overline{i}\sin 2\overline{\omega}, \quad (31)$$

$$\frac{d\delta e}{dt} = b_1\delta e + b_2\delta i + b_3\delta\omega \quad (32)$$

$$+ \frac{15fm_0}{4a_1^3 n}\overline{e}\sqrt{1-\overline{e}^2}\eta\sin 2(\chi - \overline{\Omega} + \overline{\omega}).$$

Since the coefficients b_j $(j = 1, 2, 3)$ are the partial derivatives of the right-hand side of Eq. (31) with respect to the orbital elements e, i, and ω, respectively, their order of smallness is not larger than $\sin i$ in the case under consideration $(\sin i \approx 0)$. Disregarding the products $\delta e \sin^2 i$, $\delta i \sin i$, and $\delta\omega\sin^2 i$ in Eq. (32) and assuming (to a first approximation) that

$$\overline{e} = e_0, \quad \overline{\omega} = \omega_0 + \dot{\omega}(t - t_0), \quad (33)$$
$$\overline{\Omega} = \Omega_0 + \dot{\Omega}(t - t_0)$$

on the right-hand side of this equation, where $\dot{\omega} = 2\pi/T_\omega$, $\dot{\Omega} = 2\pi/T_\Omega$, we obtain

$$\delta e(t) = \frac{15fm_0}{8a_1^3 n}\frac{e_0\sqrt{1-e_0^2}}{\dot{\omega} - \dot{\Omega}} \quad (34)$$

$$\times \eta\left[\cos 2(\omega_0 - \Omega_0 + \chi) - \cos 2(\overline{\omega} - \overline{\Omega} + \chi)\right].$$

Note that the circulation periods ω and Ω for the satellite S/2000 S1 are approximately equal to 146 and 250 years, respectively. The filled triangles in Fig. 2 indicate the dependence $e(t) = \overline{e}(t) + \delta e$ in which $\overline{e}(t)$ is determined by the general solution of the doubly averaged Hill problem (Vashkov'yak 1999) and $\delta e(t)$ is defined by formula (34).

The difference between the dependences $e(t)$ is more marked for the orbit of Saturn's fictitious satellite S/2000 S1f, which differs from the actual orbit only by one element—the inclination ($i_f = 176°$). They are shown in Fig. 3.
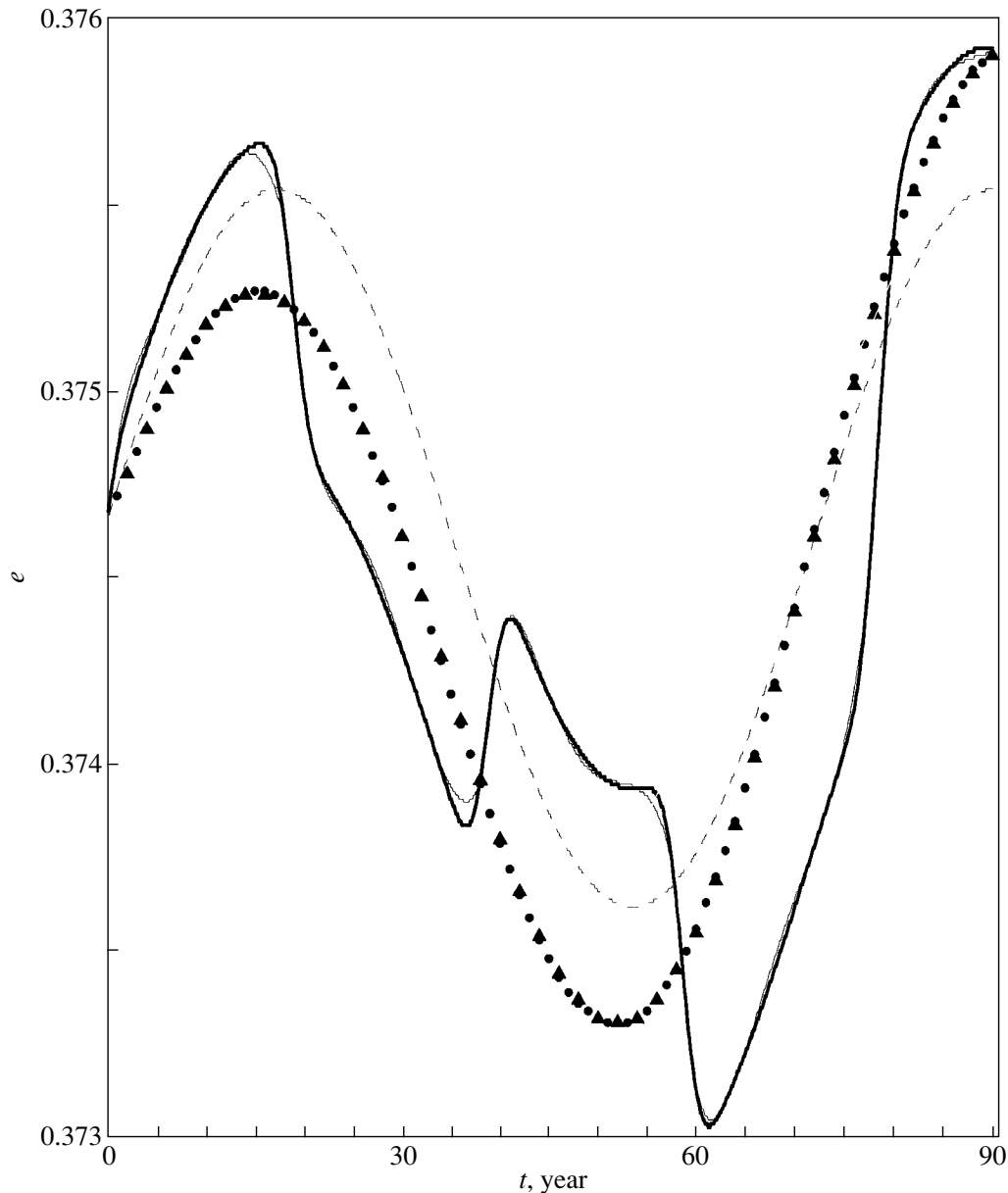
Fig. 3. The time dependence of the orbital eccentricity of the fictitious satellite S/2000 S1f ($i_\mathrm{f} = 176°$). The notation is the same as that in Fig. 2.

CONCLUSIONS

We have revealed a methodological possibility of using the wavelet approximation of the perturbing function in one specific evolution problem of celestial mechanics. In our opinion, using even a very simple burst function led to a good qualitative and a satisfactory quantitative description of the long-period orbital evolution of Saturn's most distant satellite, which is perturbed by Jupiter.

A better approximation could probably be obtained by using more complex "Mexican hat" or Morlet-type wavelets. The former contains the expression

$1 - \left(\dfrac{t - t_k}{\sigma}\right)^2$ as a cofactor to function (16). Using

the latter wavelet involves separating out the real part of the product of function (16) by $\exp[\sqrt{-1}\alpha(t - t_k)]$, where α is a parameter. Based on such more complex functions, we hope to describe not only the maximum bursts of W but also its lower-amplitude wavelike variations between the times t_k.

It seems of interest to determine whether the wavelet approximation can also be applied to strongly perturbed problems of celestial mechanics, in particular, to the problem of close encounters of comets and

asteroids with planets, as well as to the problem of the gravitational maneuvering of a spacecraft.

ACKNOWLEDGMENTS

REFERENCES

1. *Handbook of Mathematical Functions,* Ed. by M. Abramowitz and I. A. Stegun (Nauka, Moscow, 1979); Nat. Bureau Stand., Appl. Math. Ser. **55** (6) (1964).
2. C. K. Chui, *Introduction to Wavelets* (Academic Press, Boston, 1992; Mir, Moscow, 2001).
3. Y. Kozai, Astron. J. **67**, 591 (1962).
4. M. L. Lidov, Issk. Sputniki Zemli **8**, 5 (1961).
5. B. G. Marsden, MPE Circ, 2001, T06.
6. M. A. Vashkov'yak, Pis'ma Astron. Zh. **25**, 554 (1999) [Astron. Lett. **25**, 508 (1999)].

Translated by A. Dambis

An Efficient Method for Studying the Stability and Dynamics of the Rotational Motions of Celestial Bodies

A. I. Pavlov[1,2]* and A. J. Maciejewski[3]**

[1]*Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia*
[2]*Universität Mannhein, Germany*
[3]*Institute of Astronomy, Zielova Gora University, Poland*
Received November 26, 2002

Abstract—We use the alternative MEGNO (Mean Exponential Growth of Nearby Orbits) technique developed by Cincotta and Simó to study the stability of orbital–rotational motions for plane oscillations and three-dimensional rotations. We present a detailed numerical–analytical study of a rigid body in the case where the proper rotation of the body is synchronized with its orbital motion as 3 : 2 (Mercurian-type synchronism). For plane rotations, the loss of stability of the periodic solution that corresponds to a 3 : 2 resonance is shown to be soft, which should be taken into account to estimate the upper limit for the ellipticity of Mercury. In studying stable and chaotic translational–rotational motions, we point out that the MEGNO criterion can be effectively used. This criterion gives a clear picture of the resonant structures and allows the calculations to be conveniently presented in the form of the corresponding MEGNO stability maps for multidimensional systems. We developed an appropriate software package.
© 2003 MAIK "Nauka/Interperiodica".

Key words: *celestial mechanics; Solar System—planets, comets, minor bodies, and heliosphere.*

INTRODUCTION

In general, the spatial motions of the bodies in our Solar System have a complex pattern that is determined both by perturbations from other bodies and by their own internal structures. Nevertheless, many large planets and planetary satellites are characterized by a common feature or, more specifically, by a dynamically stable resonant behavior. Because of this behavior, the motions can persist for a long period. In this case, resonant couplings are observed not only between the orbital motions of planets and satellites (see, e.g., Molchanov 1973) but also between their orbital and rotational motions (see, e.g., Beletskii and Khentov 1995). Thus, the Moon is a classical example of a 1 : 1 resonance. However, because of the perturbations, the resonance ratios are never strictly fixed but are only near an exact resonance; i.e., the so-called librations are observed. Analysis of the stability conditions for quasi-stationary resonant motions allows us not only to interpret the available information on the motions of celestial bodies but also to predict some of the characteristics of the bodies themselves (see Beletskii and Khentov 1995; Beletskii 1967). In

particular, as we point out below, this analysis makes it possible to estimate the upper limit for the oblateness of a model Mercury.

The application of a reliable and sensitive technique is required to consider the range of problems pertaining to the analysis of the stability conditions for the motions of celestial bodies. This technique would allow the regular and stochastic components of the phase space of dynamical systems to be estimated rapidly and effectively.

Interestingly, the search for solutions of the celestial-mechanical equations led Poincaré to suggest that the solutions of many dynamical problems are sensitive to initial conditions. Therefore, the particular features of the orbital motions of bodies prove to be unpredictable. The salient feature of a chaotic motion is its strong sensitivity to small variations in the initial conditions. In conservative systems, chaotic trajectories uniformly fill all parts of some subspace in the phase space; i.e., they are characterized by a uniform probability density in a bounded phase space. In the course of time, nearby trajectories diverge exponentially, in contrast to regular trajectories whose divergence is only linear.[1] A quantitative estimate of

*E-mail: `apavlov@inasan.rssi.ru,pavlov@ari.`
`uni-heidelberg.de`
**E-mail: `maciejka@dede.ia.uz.zgora.pl`

---

[1]Note that such a divergence, clearly, cannot continue indefinitely in bounded space.

the divergence rate for trajectories can be represented in terms of Lyapunov indicators, which serve as a measure of the mean divergence rate of closely spaced trajectories.

The main difficulties in representing the evolution of a trajectory in phase space are related to dimensionality. Even for a system with two degrees of freedom whose phase space is four-dimensional (in the conservative case, the energy surface is three-dimensional), motion is difficult to trace, especially when we deal with the two-dimensional sheet of paper on which the image must be. One of the most valuable methods, known as the Poincaré section, which was developed by Poincaré (1971) and Birkhoff (1927), is associated with the solution of this problem. It is convenient for studying the dynamics of low-dimensional systems. For systems with a large number of degrees of freedom, such a measure of trajectory divergence as the Lyapunov indicator is still applicable.

Another method of studying the behavior of trajectories involves calculating their power spectrum (see, e.g., Tabor 1989), which is defined as the Fourier transform of the correlation function of a certain variable. In a chaotic regime, the spectrum of an irregular trajectory is much more complex than the spectrum of a regular trajectory: "islands of noise" appear around the main peaks. In general, the difference between the spectra of regular and irregular motions is very large and gives a clear idea of the dynamical systems. The FMA (Frequency Map Analysis) method developed by Dumas and Laskar (1993) and Laskar (1993) is a powerful tool that allows the frequencies associated with the Kolmogorov—Arnold—Moser tori to be accurately determined. However, in studying a motion that encompasses a sufficiently large part of the phase space, the FMA method proves to be less efficient than a direct calculation of the Lyapunov indicator. In turn, the slow convergence of the Lyapunov indicator imposes severe restrictions on computer resources. Therefore, new methods for analyzing the phase space were developed recently.

One of these new methods for studying the dynamical behavior of systems is the MEGNO indicator of chaos that was developed by Cincotta and Simó (2000). It makes it possible to analyze systems faster than with other indicators of chaos. As we show below, having a number of advantages over other techniques, this method allows the calculations to be conveniently presented in the form of the corresponding MEGNO stability maps when studying systems with a large number of degrees of freedom, in particular, when investigating the rotational dynamics of celestial bodies.

Despite the progress in computer technology and the construction of complex models for studying the stability of the Solar system (Laskar 1989), the perturbed rotational motions of celestial bodies can be described only roughly. Therefore, it seems reasonable to consider models with a gradual complication of their parameters to explain the various evolutionary—dynamical effects of the resonant motions of celestial bodies.

In Section 1, we briefly describe the MEGNO technique. In Sections 2 and 3, it is used to analyze the plane oscillations of Mercury and spatial rotations. In Section 2, we study a plane resonant rotation, i.e., a rotation where one of the principal central axes of inertia of the celestial body coincides with the normal to the Keplerian orbital plane of this body, which allows the spin—orbit interaction to be adequately studied. The Poincaré mapping method and the Floquet—Lyapunov theory are used to investigate the stability of the periodic solution of Mercury's rotation. In Conclusions, we show the distinctive features of the MEGNO indicator that make it possible to analyze the system both in the case of plane oscillations and in a more complex spatial model.

1. THE MEGNO TECHNIQUE

Below, we briefly describe the fundamentals of the MEGNO indicator based on its close relationship to the Lyapunov indicator.

As was noted in the Introduction, Lyapunov indicators are very useful in describing dynamical systems, and they make it possible to directly calculate the hyperbolicity. In order not to lose the integrity in the presentation of our material, let us first consider an autonomous system that is described by the differential equations

$$\frac{d}{dt}\mathbf{x} = \mathbf{F}(\mathbf{x}). \qquad (1)$$

We linearize the equations of motion near an arbitrary basic trajectory

$$\frac{d}{dt}\boldsymbol{\delta} = \mathbf{A}(t)\boldsymbol{\delta}, \qquad (2)$$

where $\mathbf{A}(t) = A_{ij}(t)$ is the matrix of the linearized system with the corresponding elements,

$$A_{ij}(t) = \frac{\partial F_i}{\partial x_j}(\mathbf{x}(t)).$$

The norm

$$||\boldsymbol{\delta}(t)|| = \sqrt{\sum_{i=1}^{n} \delta_i^2(t)}$$

specifies the measure of divergence of two nearby trajectories, i.e., the basic trajectory and the neighboring trajectory with the initial conditions $\mathbf{x}(t_0) + \boldsymbol{\delta}(t_0)$.

The mean rate of exponential divergence is defined as

$$\sigma := \lim_{t \to \infty} \sigma(t), \qquad (3)$$

$$\sigma(t) := \lim_{\|\boldsymbol{\delta}(t_0)\| \to 0} \left(\frac{1}{t} \right) \ln \left(\frac{\|\boldsymbol{\delta}(t)\|}{\|\boldsymbol{\delta}(t_0)\|} \right).$$

Thus, for multidimensional mappings ($\mathbf{x}$ and $\mathbf{F}$ are n-dimensional vectors), we have a set of n characteristic Lyapunov indicators that correspond to n eigenvalues of the corresponding tangent mapping. These quantities are called the characteristic Lyapunov indicators and can be ordered as

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n. \qquad (4)$$

In practice, calculating the indicators for n-dimensional fluxes reduces to calculating the largest Lyapunov indicator σ_1 from set (4), because it can be shown that, in the course of time, a small element of volume will be extended to the greatest extent in the direction that corresponds to the largest indicator, i.e., $\sigma = \lim_{t \to \infty} \sigma_1(t)$. Benettin *et al.* (1980) gave a practical scheme for calculating both the largest Lyapunov indicator and the complete spectrum σ_i.

Since σ can be analytically calculated only in very rare cases, numerical algorithms are mainly used. Their main problem is a large expenditure of computer time, because the convergence of $\sigma(t) \to \sigma$ is slow. Therefore, new approaches that allow the phase space to be studied on relatively short time scales were actively developed in recent years: the fast Lyapunov indicators (Froeschle *et al.* 1997; Froeschle and Lega 2000), the dynamical spectra (Voglis *et al.* 1999), and the FMA method mentioned in the Introduction.

The MEGNO indicator of chaos belongs to the family of fast indicators. It possesses properties that have significant advantages in studying the dynamics of planetary systems and is based on serious analytical and numerical studies.

Before writing the definition of the MEGNO indicator, we turn to the definition of the Lyapunov indicator (3), which can be represented as

$$\sigma := \lim_{t \to \infty} \frac{1}{t} \int_0^t \frac{\dot{\delta}(s)}{\delta(s)} ds = \left\langle \frac{\dot{\delta}}{\delta} \right\rangle,$$

where

$$\delta = \|\boldsymbol{\delta}\|, \qquad \dot{\delta} = \frac{d}{dt} \delta = \frac{\dot{\boldsymbol{\delta}} \cdot \boldsymbol{\delta}}{\|\boldsymbol{\delta}\|}.$$

The MEGNO indicator can then be defined as (Cincotta and Simó 2000)

$$Y(t) := \frac{2}{t} \int_0^t \frac{\dot{\delta}(s)}{\delta(s)} s \, ds; \qquad (5)$$

here, the value of $Y(t)$ is used together with its mean

$$\bar{Y}(t) := \frac{1}{t} \int_0^t Y(s) ds.$$

It can be shown (Cincotta and Simó 2000; Cincotta *et al.* 2003) that:

(A) If $\mathbf{x}(t)$ is a regular, stable solution with a linear divergence of nearby trajectories, then for a quasi-periodic solution,

$$\lim_{t \to \infty} \bar{Y}(t) = 2. \qquad (6)$$

(B) If $\mathbf{x}(t)$ is a stable, periodic solution, then

$$\lim_{t \to \infty} \bar{Y}(t) = 0. \qquad (7)$$

(C) For a chaotic behavior of $\mathbf{x}(t)$,

$$\bar{Y}(t) \approx \frac{\sigma}{2} t \quad \text{for} \quad t \to \infty,$$

where σ is the largest Lyapunov indicator for $\mathbf{x}(t)$.

If $\bar{Y}(t)$ approaches a certain fixed value that differs from 2, then the closely spaced trajectories diverge as a power law.

The MEGNO time evolution can be represented in a universal form for any type of motion. This important property of the MEGNO indicator significantly distinguishes it from the Lyapunov indicator. Indeed, representations (A)–(C) can be combined, because the asymptotic behavior of $\bar{Y}(t)$ is expressed as

$$\bar{Y}(t) \simeq at + d,$$

where $a = 0$, $d \simeq 0$ or 2 for stable periodic and quasi-periodic motions or $a = \sigma/2$, $d \simeq 0$ for an irregular stochastic motion.

This property can be used to effectively estimate σ, which can be determined by linearly approximating $\bar{Y}(t)$. This method for calculating the Lyapunov indicator requires much less computer time and is more efficient and sensitive to the structure of the phase space, because $\bar{Y}(t)$ contains the complete dynamical information on the system obtained over the entire integration time. These properties make the MEGNO indicator a useful tool for studying fine structures in phase space: they reveal resonances and allow their widths to be estimated. In this paper, we use some of these properties to study spin–orbit interactions.

The following estimates clearly show that the MEGNO indicator converges faster than the Lyapunov indicator ($t \to \infty$):

$$\frac{Y(t)}{t} \simeq \frac{2}{t} \qquad (8)$$

for a regular solution and

$$\frac{Y(t)}{t} \simeq \sigma \qquad (9)$$

for chaotic behavior. Thus, for regular motion, $Y(t)/t$ converges to zero faster than σ (whose convergence rate is $(1/t)\ln t$), while for chaotic motion, the rate of convergence to a certain nonzero value is approximately the same.

As was noted by Goździewski *et al.* (2001), the algorithm for calculating the MEGNO indicator is simple: it is necessary to write the equations of motion of the system (1) together with its variational equations (2) and then add two more equations:

$$\frac{d}{dt}y = \frac{\dot{\boldsymbol{\delta}} \cdot \boldsymbol{\delta}}{\boldsymbol{\delta} \cdot \boldsymbol{\delta}}t, \qquad \frac{d}{dt}w = 2\frac{y}{t};$$

then, $Y(t) = 2y(t)/t$ and $\bar{Y}(t) = w(t)/2$.

The dynamics of the motion proves to be hidden when the Lyapunov indicator is calculated, because integrating the system over a long but finite period ($\sim 10^6 - 10^8$ characteristic time scales of the system), we obtain only one number. When performing numerical calculations of the dynamics of the system over a finite period whose net result is determined by a finite nonzero value of σ, it is difficult and, in several cases, virtually impossible to confirm that the Lyapunov indicator points to a regular case. An example of such a situation when calculating σ can be found in the paper of Sussman and Wisdom (1988) who analyzed the orbital stability of Pluto. After more than 10^8 years of quasi-regular behavior, the system exhibited a nonzero σ with a convergence rate of $10^{-7.3}$ yr^{-1}. Thus, the Lyapunov indicator does not always allow regular and chaotic motions to be classified reliably and rapidly.

The MEGNO technique is a powerful tool for identifying the regular and stochastic components of the system, because the chaotic effects in the behavior of the solutions to the variational equations are enhanced (see definition 5). Thus, we can analyze the system, on average, two to four orders of magnitude faster than with the Lyapunov indicator.[2] Such an efficiency is an indispensable factor in studying the global dynamics, which requires analyzing a large set of initial data. This method for analyzing planetary systems was first implemented by Goździewski *et al.* (2001) and Goździewski and Maciejewski (2001).

We will not consider in detail the convergence properties of the MEGNO indicator, which are discussed both in the series of papers of its authors

Fig. 1. The geometry of model plane resonant rotations: φ is the angle between the principal central axis of inertia of the body and the radius vector R of its center of mass, ν is the true anomaly, T is the orbital period of Mercury, and C.P. is the central planet (Sun).

(Cincotta and Simó 2000; Cincotta *et al.* 2003) and in the paper of Goździewski *et al.* (2001), but focus our attention on the possibility of using MEGNO to study the fine resonant structures in the rotation of a celestial body in 3 : 2 synchronism as an example.

For these purposes, we developed a software package in the IDL interactive programming language using the ODEX numerical integrator (Hairer and Wanner 1995) to solve systems of ordinary differential equations. This software package has a sophisticated graphical user interface with a branched menu, which allows us to change system parameters and form images interactively.

2. ANALYSIS OF THE PLANE RESONANT ROTATIONS OF A BODY

The model resonant rotation of a celestial body in an unperturbed Keplerian orbit about one of its principal axes of inertia perpendicular to the orbital plane allows the resonant spin−orbit couplings to be adequately analyzed.[3] Figure 1 shows a schematic view of the Sun−Mercury system. We see that the Mercurian type of motion traces the pericentric radius vector of the orbit.

Let us write the equation for the plane rotation of a celestial body (Mercury) in the gravitational field of a mass point (Sun) by concentrating the entire

---

[2] Note that this estimate for the speed of the MEGNO indicator does not follow from the estimates of the convergence rate (8), (9), but it was estimated by Cincotta and Simó (2000) and Goździewski *et al.* (2001) numerically.
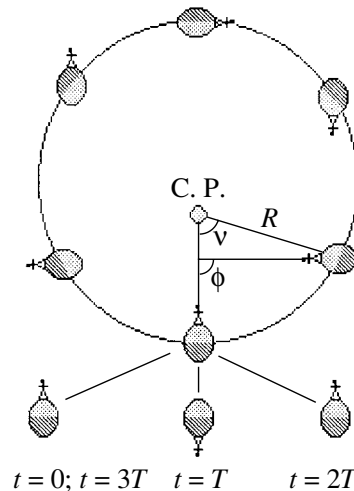
[3] Various aspects of the resonant rotations of celestial bodies are considered in the book by Beletskii and Khentov (1995). Note that Celleti and Falcolini (1992) and Celleti and Chierchia (2000) investigated the dynamics and stability of the spin−orbit interactions in terms of perturbation theory.
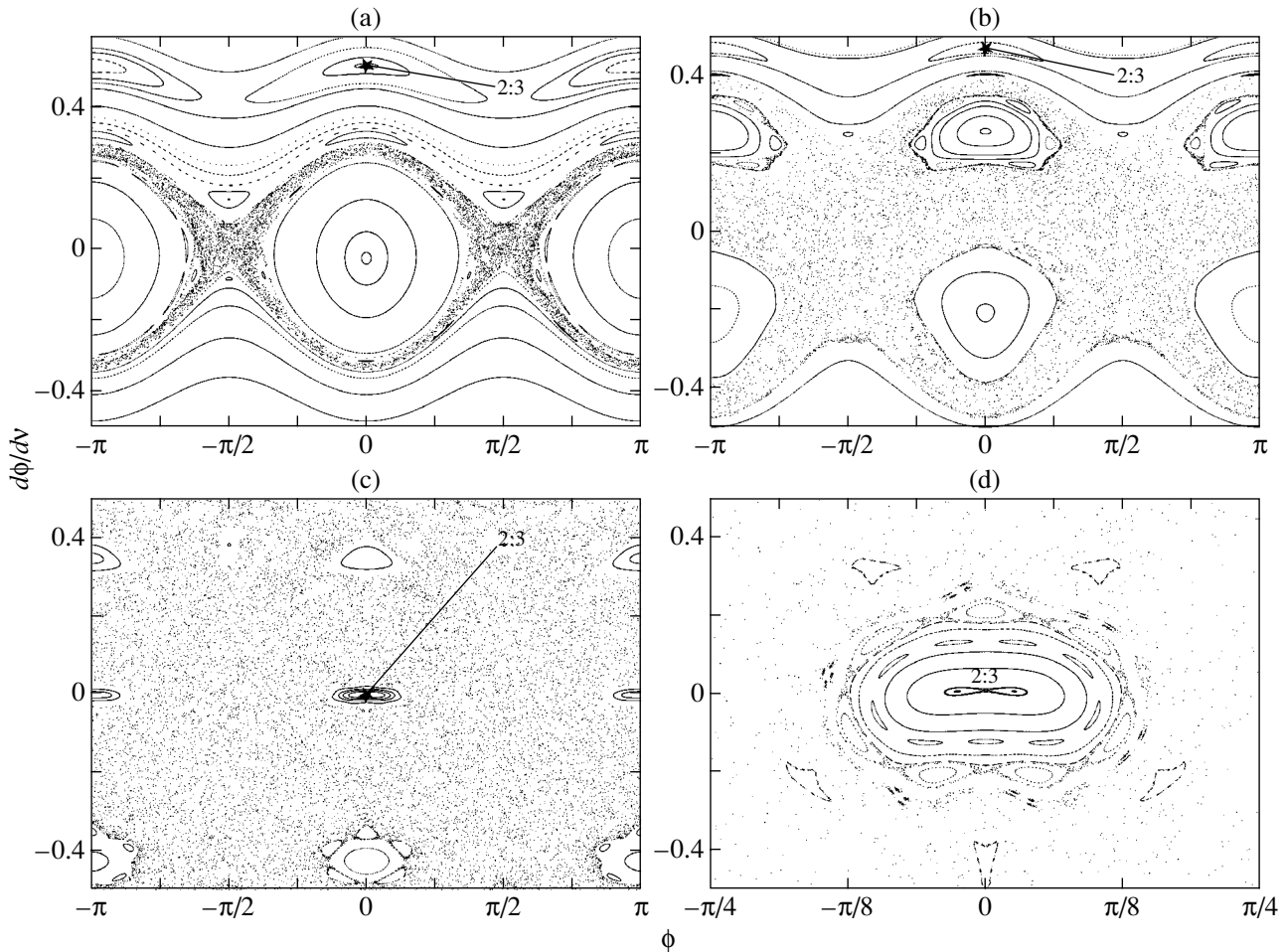
Fig. 2. The Poincaré sections. The stability of the solution as a function of the ellipticity parameter of the body d and the orbital eccentricity e: (a) $d = 0.1$, $e = 0.001$; (b) $d = 0.1$, $e = 0.1$; (c) $d = 0.195$, $e = 0.2056$; and (d) $d = 0.197$, $e = 0.2056$.

mass of the Sun at its center of mass and analyze a Mercurian-type resonant rotation following Beletskii and Khentov (1995). In this case, Mercury is treated as an absolutely rigid body. Thus, we leave aside the questions related to the change of its shape, which corresponds to the neglect of tidal effects. In turn, tidal effects introduce dissipation into the spin–orbit interaction and, in general, can affect the dynamical evolution of the system on long time scales.

The motion under consideration can be described by the second-order differential equation (Beletskii 1965)

$$(1 + e\cos\nu)\frac{d^2\phi}{d\nu^2} - 2e\sin\nu\frac{d\phi}{d\nu} \qquad (10)$$
$$+ d\sin\phi\cos\phi = 2e\sin\nu,$$

where ϕ is the angle between the principal central axis of inertia of a body that lies in the plane of an elliptical orbit of eccentricity e and the radius vector R of its center of mass and ν is the true anomaly. The

ellipticity parameter of the body is defined as

$$d = 3(B - A)/C;$$

A, B, and C denote the principal central moments of inertia of the celestial body.

After the standard reduction of the second-order differential equation to a system of first-order ordinary differential equations (FODEs), we performed our calculations numerically by using the ODEX integrator. The ODEX integrator, which was developed by Hairer (Hairer and Wanner 1995), includes the Bulirsch–Ster extrapolation method (see, e.g., Press *et al.* 1992), and, in addition to computational accuracy, it has a high speed.

The Poincaré Section. The Soft Loss of Stability

We present the results of our calculations after each complete period by taking into account the fact that, for the resonance under consideration, the angle ϕ must change by 2π over two orbital revolutions.
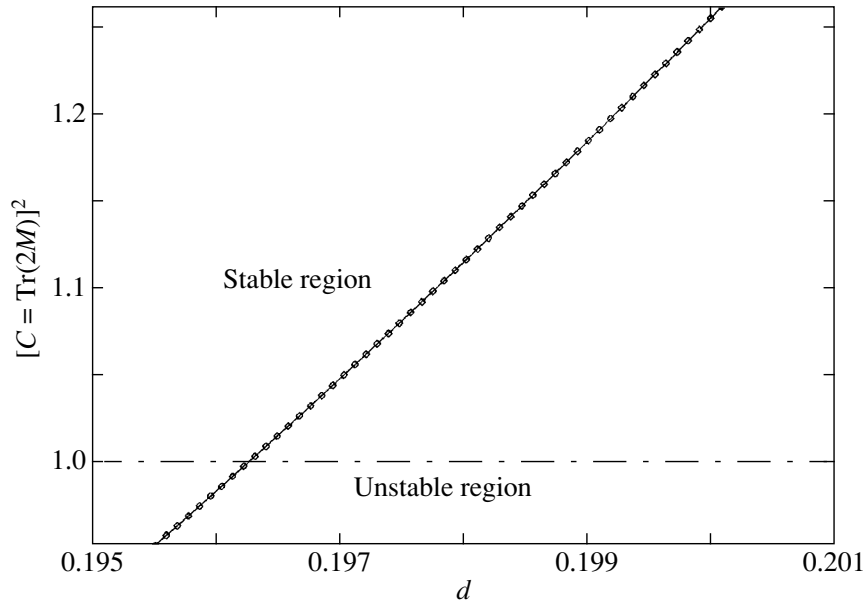
Fig. 3. The stability of the solution as a function of the body's ellipticity parameter d.

Thus, we construct a Poincaré section that is used to simplify the observed pattern of complex time regimes and pass from the flux to a lower-dimension mapping.

Our numerical calculations for various values of e and d presented in Fig. 2 show a strong dependence of the pattern of motion on the system's parameters.[4] We see that the region around the resonance in $3:2$ synchronism (marked by the asterisk) changes, becoming much smaller with increasing (e, d). The chaotization of the system's motion is a manifestation of the instability of the phase trajectories that arises under certain conditions. Nevertheless, the region of quasi-stationary motions is still large enough even for a large ellipticity parameter of the planet.

Let us now analyze the stability of the system's periodic solution (asterisk) as a function of the ellipticity parameter of the celestial body by fixing the eccentricity that corresponds to the Mercurian orbit (Figs. 2c, 2d). For this purpose, we consider the periodic solution of our system with a period T ($\mathbf{x}(t + T) = \mathbf{x}(t)$).

As a rule, periodic solutions are analytically studied in a general form. The Poincaré section is most commonly constructed, using which studying the properties of a periodic trajectory reduces to analyzing the properties of a fixed point of some mapping—the Poincaré mapping. In studying the stability of this fixed point, we linearize the equations of motion

---

[4]Figures 2c and 2d present the calculations that correspond to the eccentricity of Mercury.

in its vicinity (see Eq. (2)). Since the fundamental matrix $\Phi(t)$ satisfies the equation $\dot{\Phi} = A(t)\Phi$,

$$\dot{\Phi}(t + T) = A(t + T)\Phi(t + T) = A(t)\Phi(t + T);$$

i.e., $\dot{\Phi}(t + T)$ is another fundamental matrix. Since any solution can be expressed as a combination of the fundamental solutions,

$$\dot{\Phi}(t + T) = \Phi(t)M,$$

where M is a constant matrix called a monodromy matrix. Thus, the Poincaré section and the Floquet–Lyapunov theory (see, e.g., Kozlov 1996) allow the problem of the stability of a periodic solution to be reduced to the problem of the stability of a fixed mapping point, which is very useful in investigating the bifurcations of the periodic regimes. Integrating the variational equation (2) over the complete period yields a monodromy matrix whose eigenvalues μ_i (Floquet multipliers) control the growth rate of the perturbation.

In our case, the characteristic equation for the linear second-order differential equation with a 2π period coefficient can be written as (Malkin 1966)

$$\lambda^2 - 2C\lambda + 1 = 0;$$

the coefficient C is defined as

$$C = \frac{1}{2}\mathrm{Tr}M,$$

where $\mathrm{Tr}M = [y_1(2\pi) + \frac{dy_2}{d\nu}(2\pi)]$ is the trace of the monodromy matrix. In this case, the functions $y_1(\nu)$ and $y_2(\nu)$ are the fundamental system of solutions for
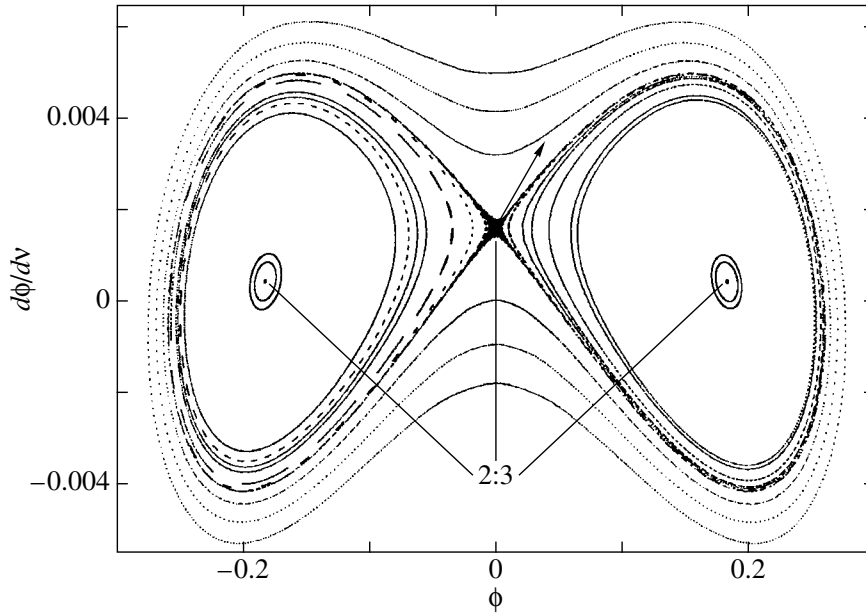
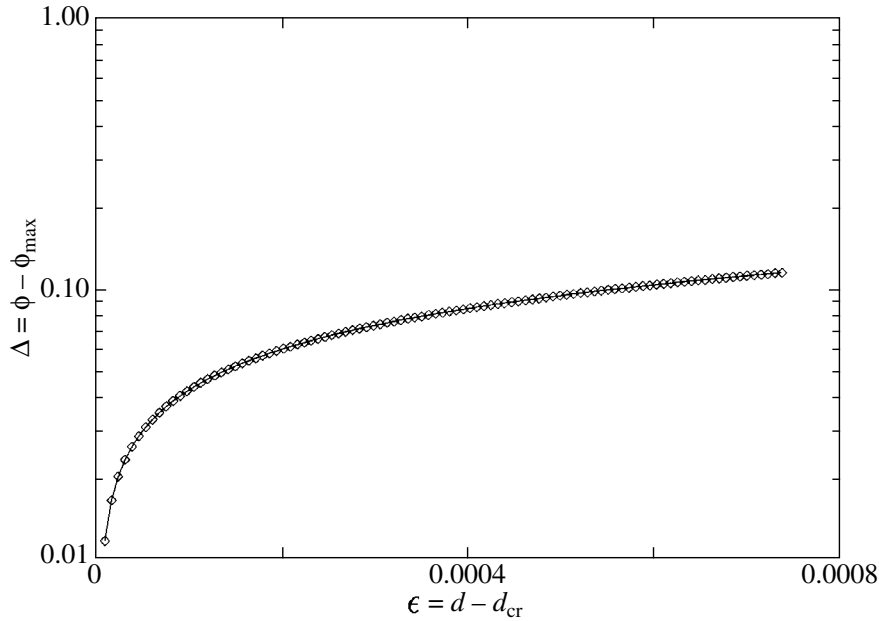**Fig. 4.** The Poincaré section: the bifurcation of the periodic solution ($d = 0.2$).

Fig. 5. The maximum deviation of the separatrix in Φ from the unstable periodic solution as a function of the body's ellipticity parameter d.

the variational equation that was normalized at the initial point $\nu = 0$, i.e., satisfies the initial conditions

$$y_1(0) = 1, \quad \frac{dy_1}{d\nu}(0) = 0,$$

$$y_2(0) = 0, \quad \frac{dy_2}{d\nu}(0) = 1.$$

The periodic solution of Eq. (10) is stable in the linear approximation if $C^2 < 1$ and unstable if $C^2 > 1$. The condition $C^2 = 1$ specifies the boundary of the stability region for the periodic solution. In our calculations, we used the method of Brent (1973) both to find the periodic solution (by scanning the Poincaré section) and to search for the boundary of the stability region $C^2 = 1$, which, together with the ODEX integrator, allowed us to achieve a result with a computational accuracy of $\sim 10^{-16} - 10^{-14}$.
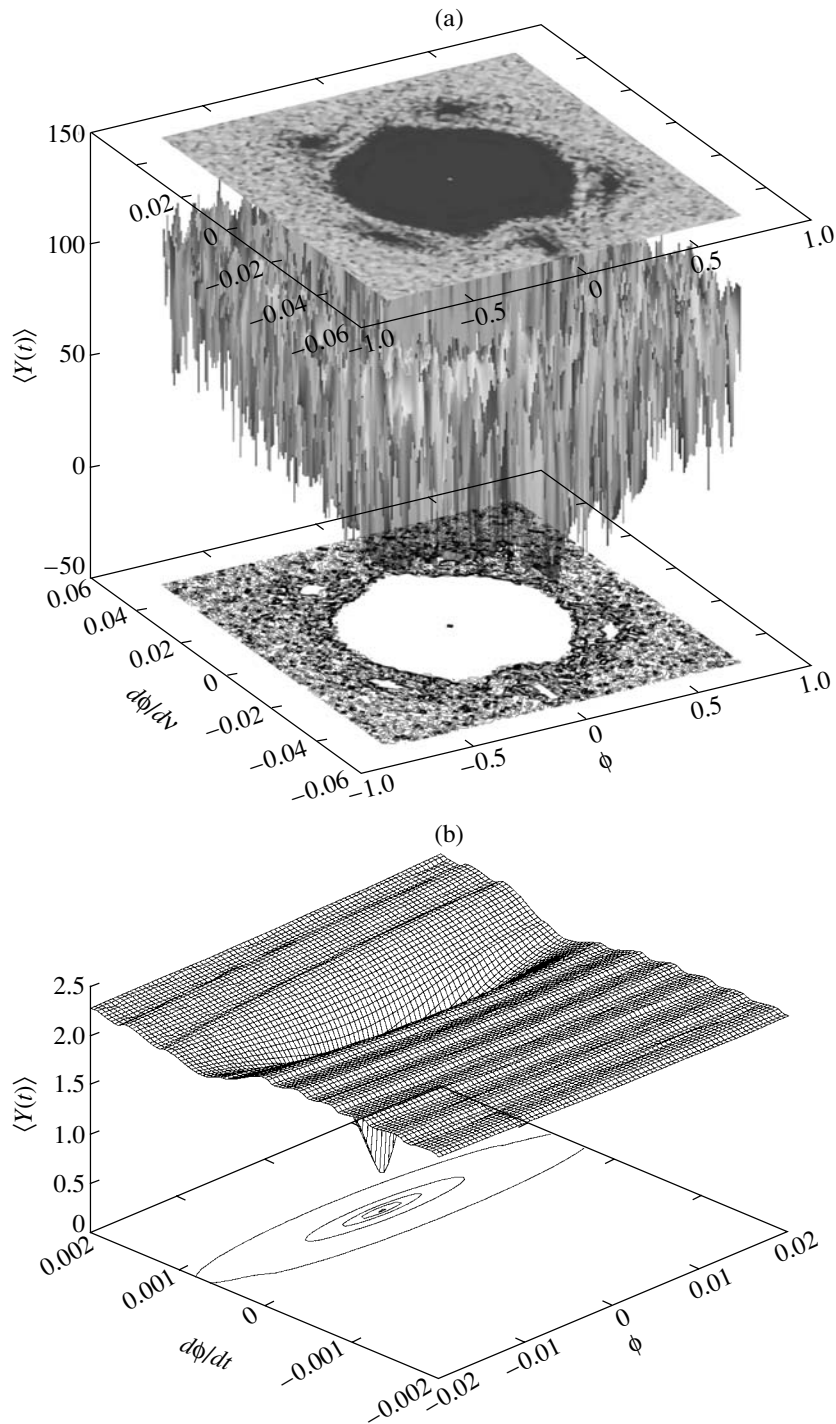
Fig. 6. (a) The MEGNO indicator for a set of initial conditions ϕ and $\dot{\phi}$ near the stable periodic solution for a 3 : 2 resonance with its contour mapping; (b) direct evidence of the periodic solution for $\bar{Y}(t) \simeq 0$.

Since Brent's method includes the bisection method and quadratic interpolation, the algorithm rapidly and assuredly converge with a prespecified accuracy.

The results of our analysis of the system's stability as a function of the ellipticity parameter are presented in Fig. 3, in which the boundary between the stable and unstable periodic solutions is indicated by the dash−dotted line.

As was noted above, the periodic solution of Eq. (10) under consideration may, in a sense, be considered as modeling the observed resonant rotation of Mercury, and the results obtained can be
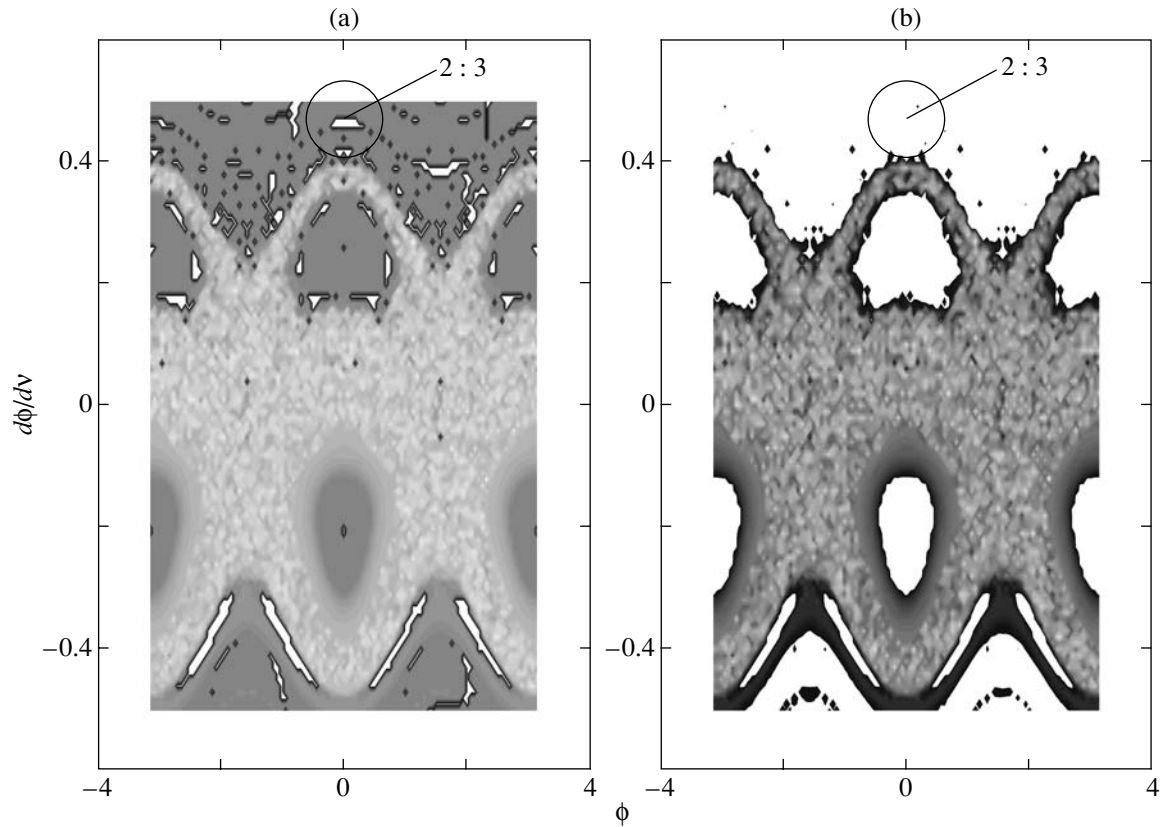
Fig. 7. The system's stability maps obtained by using (a) MEGNO and (b) the maximum Lyapunov indicator; the parameters d and e correspond to Fig. 2b.

used to estimate the upper limit for the ellipticity of Mercury. Our estimate of the critical ellipticity parameter ($d = 0.19626$) is slightly larger than that found by Beletskii and Khentov (1995) ($d = 0.1961$).[5]

Figures 2a–2c and Figs. 2d and 4 present the results of our calculations for sets of the stable and unstable periodic solutions of the $3 : 2$ resonance under consideration. The Poincaré section in Fig. 2d and, more clearly, in Fig. 4 shows the bifurcation of the periodic solution: starting from $d = 0.19626$, we obtain three periodic solutions that correspond to $3 : 2$ synchronism, two of which are stable and one is unstable. The so-called soft loss of stability is observed. In other words, when the periodic solution under consideration becomes unstable, the phase curves with initial conditions close to it remain near the periodic solution in question.

In this case, according to Treshchev (1992), the maximum size of the separatrix for a deviation from the critical parameter d_{cr} must increase as

$$\Delta \sim K\sqrt{\epsilon}, \quad \text{where} \quad \epsilon = d - d_{\mathrm{cr}}.$$

---

[5]Note that this discrepancy in the estimated ellipticity parameters probably stems from the fact that we used more accurate numerical algorithms.

The coefficient $K = 4.26$ for the resonance under consideration was determined by fitting the numerical calculations shown in Fig. 5 in the form of rhombuses. In calculating the maximum deviation from the periodic solution, we determined one of the eigenvectors of the monodromy matrix (see the vector in Fig. 4), whose direction specified the subsequent integration along the separatrix.

Thus, the periodic solution becomes unstable at $d_{\mathrm{cr}} = 0.19626$, which may serve as an upper limit on the ellipticity parameter for Mercurian-type motion. However, near d_{cr}, the instability of this periodic solution does not imply the catastrophic chaotization of the system and may be considered as a "fuzziness" of the body's critical ellipticity parameter d_{cr}. Analysis of the system's behavior as the ellipticity parameter increases further is indicative of a large deviation from the periodic solution along the separatrix.

MEGNO Maps

Let us now present the results of our stability analysis for the system under consideration by using the MEGNO indicator. In contrast to the Poincaré section, the calculations should be presented in three-dimensional space (see Fig. 6a), because now we
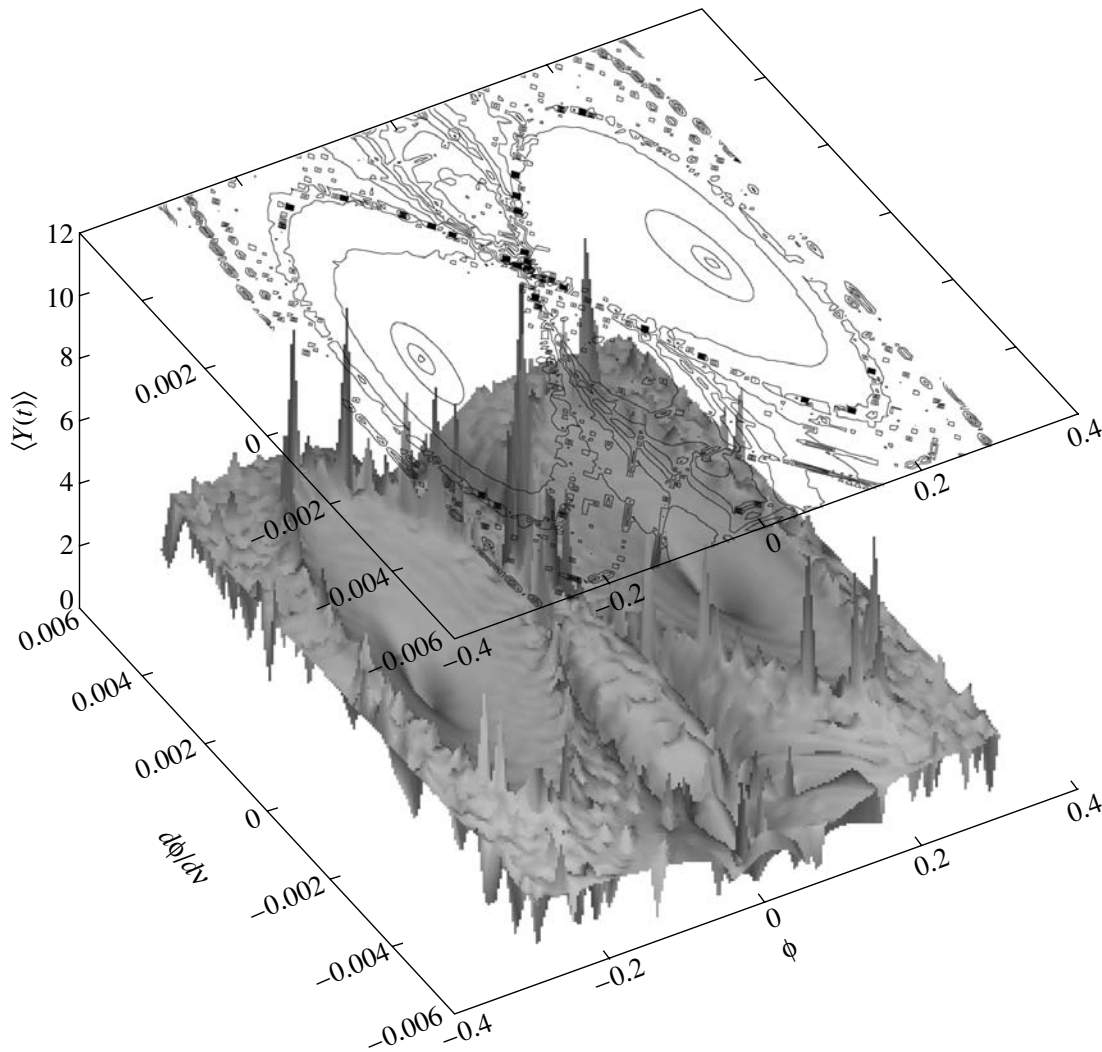
Fig. 8. The MEGNO indicator: the bifurcation of the periodic solution (an analogue of Fig. 4). In addition to one unstable periodic solution, two stable solutions are observed.

are deducing not the system's solution but the value of $\bar{Y}(t)$ that shows the stability of the system under consideration for the specified initial conditions ϕ and $\dot{\phi}$. The complexity of such a presentation becomes clearer when using a contour mapping on the $(\phi, \dot{\phi})$ plane, which either can be included as part of the figure (the upper and lower parts of Fig. 6a) or can be presented as a separate figure (Fig. 7).

Certainly, using the MEGNO indicator for plane oscillations is somewhat artificial, because complete information can be obtained with the Poincaré section (Fig. 2). However, the system in question, whose dynamical properties can be obtained from the Poincaré section, allows us to clearly show the distinctive features of the MEGNO indicator.

The MEGNO mapping shown in Fig. 7a reveals the fine structures of the phase space and the pe-

riodic solutions, which is a remarkable property of the MEGNO indicator. Indeed, according to (B), $\bar{Y}(t) \simeq 0$ for a stable periodic solution, while for a quasi-periodic solution, it tends to 2 (A). Figure 6b, which is an enlargement of the 3 : 2 resonance region, clearly shows this feature of the MEGNO indicator. The presentation of the mapping using the Lyapunov indicator (Fig. 7b) gives a similar picture and hyperbolicity estimate but does not explicitly indicate the periodic solutions. Note that neither of the methods reveal any unstable periodic solutions. At the same time, the characteristics of the regions of chaotic motions can be determined using the MEGNO indicator much more efficiently, which allows the mappings to be rapidly constructed when studying the complex structure of the phase space (Fig. 8). More specifically, the MEGNO efficiency, together with its dis-
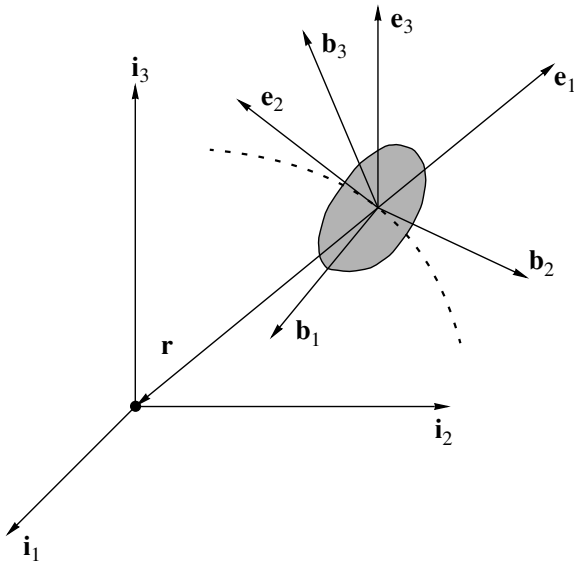
Fig. 9. The geometry of the model spatial rotations of a rigid body: $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ is an inertial frame of reference; $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ is the frame of reference associated with the body; $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is the orbital frame of reference; and $\boldsymbol{\gamma} = \mathbf{e}_1 = -\mathbf{r}/r$.

tinctive features discussed above, is an indispensable factor in studying multidimensional systems where the Poincaré section cannot be constructed and a large set of initial conditions should be considered.

3. THE SPATIAL ROTATIONS OF A CELESTIAL BODY IN AN ELLIPTICAL ORBIT

We now turn to a much more complex model to study the spatial rotations of a celestial body with an arbitrary central ellipsoid of inertia. Let the orbit of the body lie in the $\mathbf{i}_1\mathbf{i}_2$ plane of the inertial frame of reference $I(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ (see Fig. 9). Denote the corresponding unit vectors of the orbital frame of reference O by $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Then, assuming that

$$\boldsymbol{\gamma} = \mathbf{e}_1 = -\mathbf{r}/r,$$

we can write the equations of motion as (Beletskii 1965)

$$\frac{d}{dt}\mathbf{M} = \mathbf{M} \times \boldsymbol{\Omega} + 3S\boldsymbol{\gamma} \times \mathbf{I}\boldsymbol{\gamma},$$

$$\frac{d}{dt}\boldsymbol{\gamma} = \boldsymbol{\gamma} \times (\boldsymbol{\Omega} - \omega\mathbf{n}),$$

$$\frac{d}{dt}\mathbf{n} = \mathbf{n} \times \boldsymbol{\Omega}, \quad \boldsymbol{\Omega} = \mathbf{I}^{-1}\mathbf{M}, \qquad (11)$$

where $\boldsymbol{\Omega}$ and $\mathbf{M} = \mathbf{I}\boldsymbol{\Omega}$ are the angular velocity and angular momentum of the rigid body, respectively; $\mathbf{n}$

is a unit vector perpendicular to the orbital plane; and $\mathbf{I}$ is the inertia tensor of the body; i.e.,

$$\mathbf{I} = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}.$$

In turn, ω and S are expressed as

$$\omega = \frac{d\nu}{dt} = \frac{\sqrt{\mu a(1 - e^2)}}{R^2}, \qquad S = \mu/R^3,$$

where ν is the true anomaly; μ is the gravitational parameter of the Sun; and e and a are the orbital eccentricity and semimajor axis for Mercury, respectively. R is defined as

$$R = a(1 - e\cos E),$$

where E is the eccentric anomaly calculated from the Kepler equation

$$E - e\sin E = \sqrt{\mu/a^3}(t - \tau),$$

where τ is the time of periastron passage.

System (11) has three geometrical integrals,

$$h_1 = (\boldsymbol{\gamma}, \boldsymbol{\gamma}), \quad h_2 = (\mathbf{n}, \mathbf{n}), \quad h_3 = (\boldsymbol{\gamma}, \mathbf{n}),$$

and, for a circular orbit ($\omega = \text{const}$, $S = \omega^2$), has the energy integral

$$H = H(\mathbf{M}, \boldsymbol{\gamma}, \mathbf{n}) = \frac{1}{2}(\mathbf{M}, \mathbf{I}^{-1}\mathbf{M})$$

$$- \omega(\mathbf{M}, \mathbf{n}) + \frac{3}{2}\omega^2(\boldsymbol{\gamma}, \mathbf{I}\boldsymbol{\gamma}).$$

The classical description of the position of a body in six-dimensional phase space is based on the parametrization of its representation by three Euler angles (ϕ, θ, ψ) and the corresponding angular velocities. Having fixed four of the six variables under consideration, we can study the structure of this phase space by varying the two remaining variables in a certain range. Thus, we obtain a stability map of the system under consideration by scanning the domain of the two chosen variables and calculating the MEGNO indicator at each point of the plane. A set of such maps gives a comprehensive idea of the system's stability.

Figures 10 and 11 present such stability maps for the spatial rotation of a model Mercury for nutation angles of 1° and 10°, respectively.

In exactly the same way, we construct such MEGNO maps by fixing ϕ and $\dot{\phi}$ and varying the angle θ and the body's ellipticity parameter d. In Fig. 12, we analyze the system's stability as a function of the angle $\theta \in [-\pi, \pi]$ and the ellipticity parameter $d \in [0.1955, 0.2]$ for ϕ and $\dot{\phi}$ corresponding to the
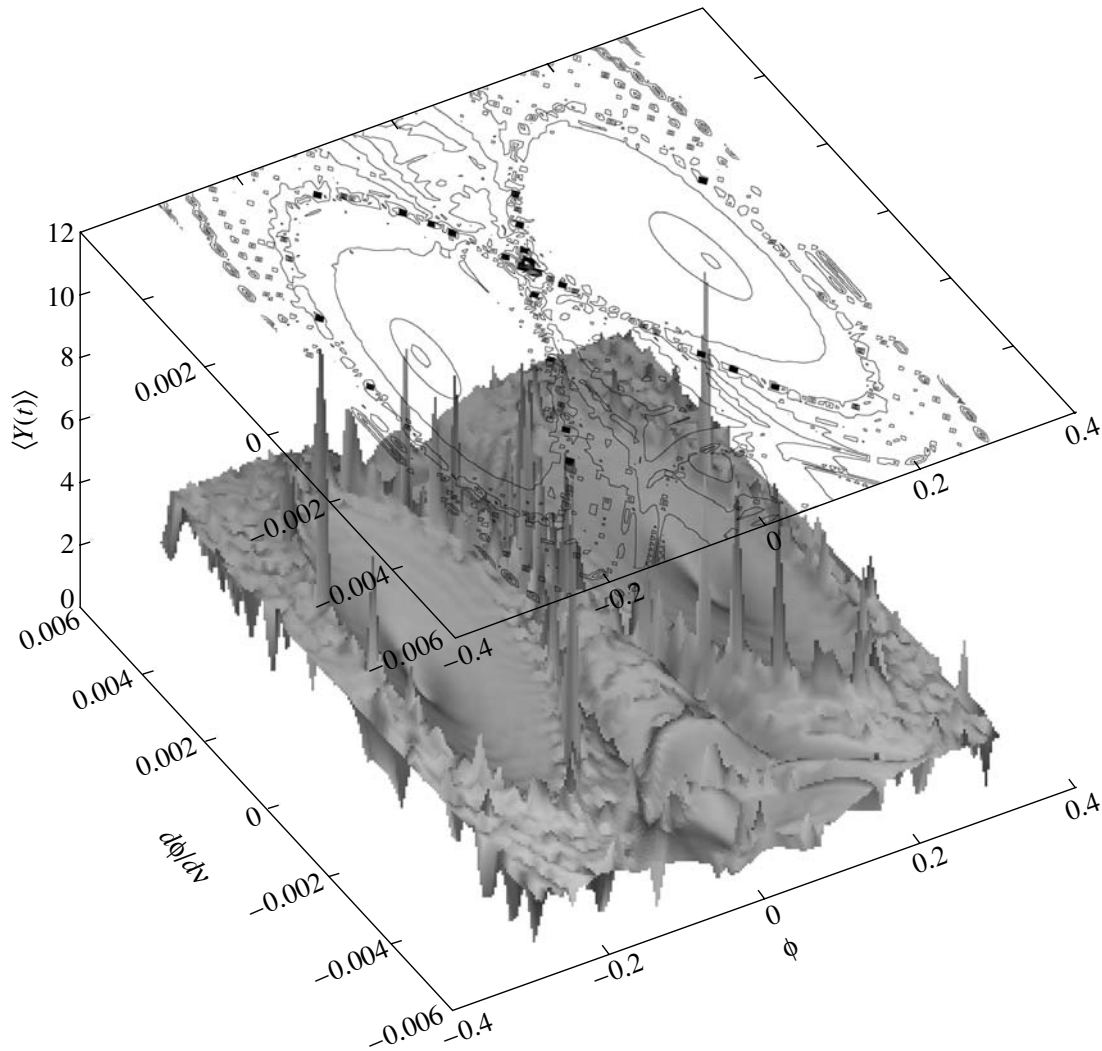
Fig. 10. The spatial model: the MEGNO indicator for a set of initial conditions ϕ and $\dot{\phi}$ (see Fig. 4) at $\theta = 1°$.

periodic solution in the case of plane oscillations ($\theta = 0$; see the preceding section). This picture allows us to determine the instability regions for the system's solution, which becomes appreciably more complex than that for plane oscillations. The auxiliary two-dimensional plots allow the behavior of the system to be traced for a certain set of initial conditions:

− the left curve, the projection of the three-dimensional pattern at $\theta = 0°$, shows the transition from the region of the stable periodic solution to the region of the unstable solution at $d = 0.19626$ (we see the jump in the indicator);

− the other two curves, which were brought to the foreground of the image for clarity, show the change in the system's stability with θ for the stable ($d = 0.1955$, solid line) and unstable ($d = 0.2$, dash−dotted line) periodic solutions in the case of Mercury's plane oscillations.

We see from the figures that, making a cut in the multidimensional phase space, we can conveniently study the behavior of the system using MEGNO maps.

We will not consider the possibility of refining the two-sided estimate of Mercury's ellipticity. We only note that Beletskii and Khentov justified and developed the fundamental hypothesis that Cassini's laws, which describe the rotation of the Moon, are generalized and are used to describe the resonant rotations of other celestial bodies. Based on these laws and observational data, Beletskii and Khentov (1995) obtained two-sided estimates for the ellipticity of Mercury.

CONCLUSIONS

The main goal of this paper is to show that the MEGNO indicator can be used to effectively study the behavior of regular and chaotic trajectories both
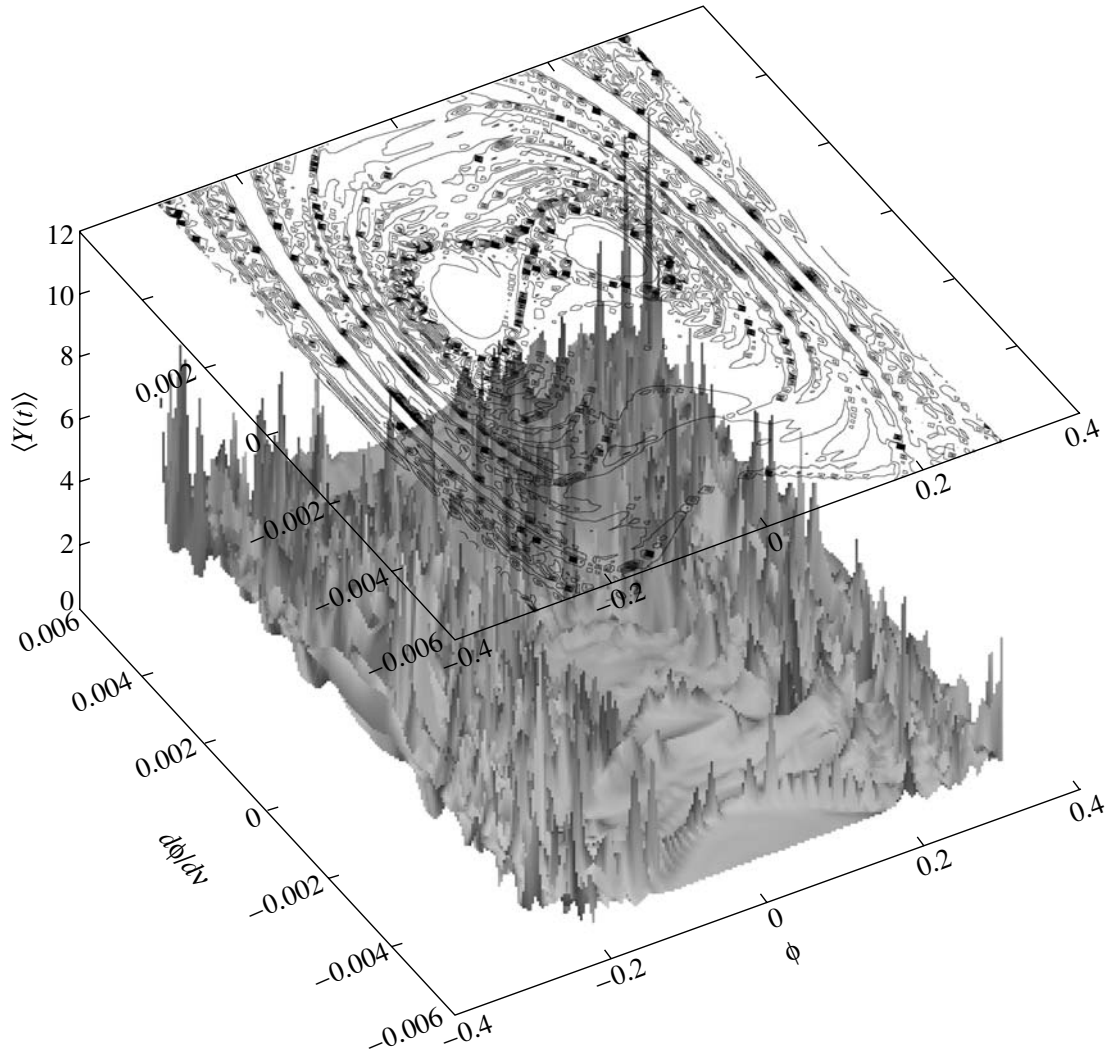
Fig. 11. The spatial model: the MEGNO indicator for a set of initial conditions ϕ and $\dot{\phi}$ (see Fig. 4) at $\theta = 10°$.

for systems with a small number of degrees of freedom (plane oscillations) and for dynamical systems with a large number of degrees of freedom (spatial rotations).

The MEGNO indicator allows the regular and stochastic components of the system to be reliably identified, because the chaotic effects in the behavior of the solutions to the variational equations are enhanced. This makes it possible not only to analyze the stability of the system faster than with other indicators of chaos but also to obtain additional information on the stability of resonant structures. We showed that this method allows the calculations to be conveniently presented in the form of MEGNO stability maps for multidimensional systems and, in combination with other methods, is an efficient tool for studying the stability of spin–orbit couplings.

We investigated the stability, to a first approxi-

mation, of the periodic solution for a Mercury-type system (3 : 2 resonance). We showed that the loss of stability is accompanied by the bifurcation of the periodic solution: when the ellipticity parameter reaches its critical value, we observe the appearance of two new periodic solutions close to the solution being studied, which, in turn, shows the soft loss of stability. We refined the upper limit of the ellipticity parameter for a model Mercury in the case of plane oscillations ($d_{\mathrm{cr}} = 0.19626$) and derived the dependence of the region of soft loss of stability on the ellipticity parameter of the body. We showed that the instability of the given periodic solution near d_{cr} does not imply a catastrophic chaotization of the system and can be treated as a "fuzziness" of the critical ellipticity parameter of the body.
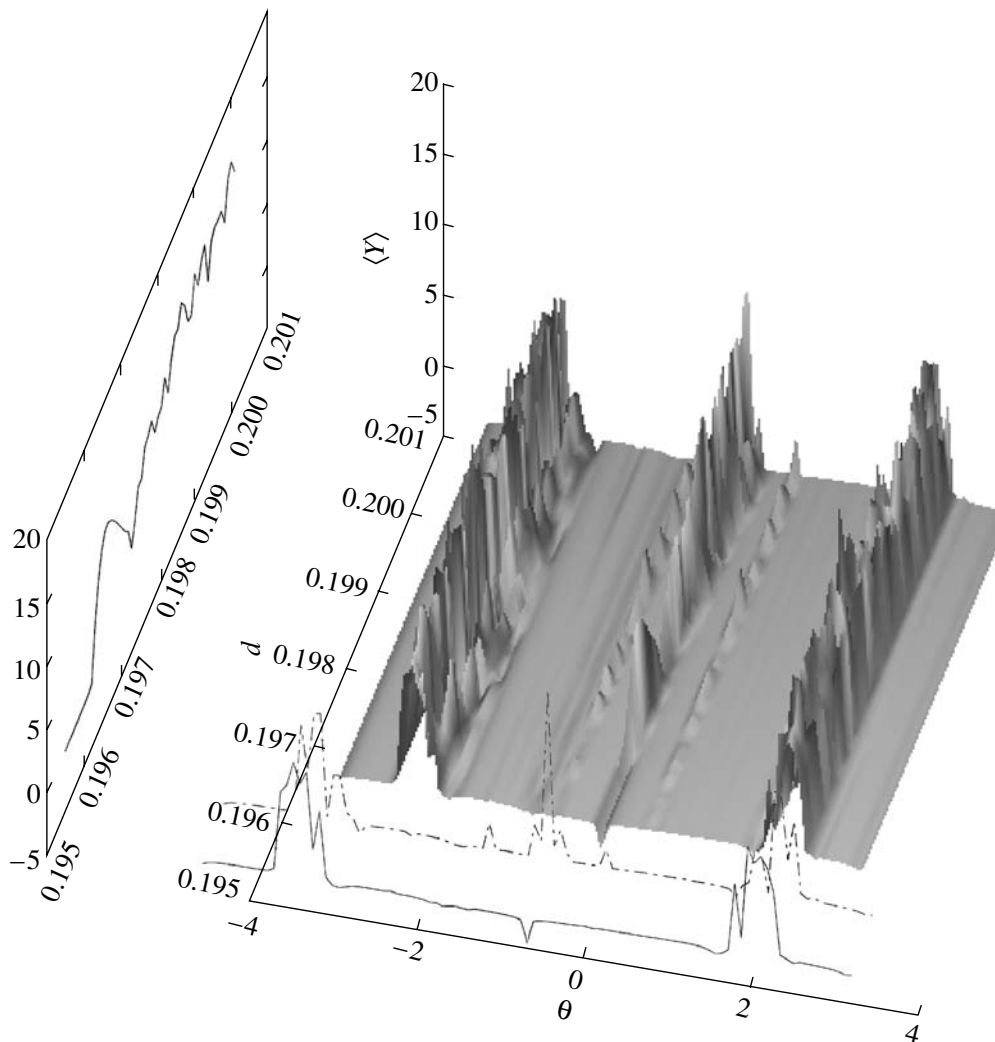
Fig. 12. The spatial model: the MEGNO indicator for $\theta \in [-\pi, \pi]$, and the ellipticity parameter $d \in [0.1955, 0.2]$.

REFERENCES

1. V. V. Beletskii, *The Motion of an Artificial Satellite about the Center of Mass* (Nauka, Moscow, 1965).
2. V. V. Beletskii, Prikl. Mat. Mekh. **31**, 1104 (1967).
3. V. V. Beletskii and A. A. Khentov, *Resonant Rotations of Celestial Bodies* (Nizh. Gumanitar. Tsentr, Nizhnii Novgorod, 1995).
4. G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica, No. 3, 9 (1980).
5. G. D. Birkhoff, *Dynamical Systems* (Am. Math. Soc, New York, 1927; GTTI, Moscow, 1941).
6. R. P. Brent, *Algorithms for Minimization without Derivities* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
7. A. Celletti and L. Chierchia, Celest. Mechan. Dynam. Astron. **76**, 229 (2000).
8. A. Celletti and C. Falcolini, Celest. Mechan. Dynam. Astron. **53**, 113 (1992).
9. P. M. Cincotta, C. M. Giordano, and C. Simó, Physica D (2003) (in press).
10. P. M. Cincotta and C. Simó, Astrophys. J., Suppl. Ser. **147**, 205 (2000).
11. H. S. Dumas and J. Laskar, Phys. Rev. Lett. **70**, 2975 (1993).
12. C. Froeschlé and E. Lega, *Am. Astron. Soc., Div. Planet. Sci. Meeting* **32**, 6001 (2000).
13. C. Froeschle, E. Lega, and R. Gonczi, Celest. Mechan. Dynam. Astron. **67**, 41 (1997).
14. K. Goździewski and A. J. Maciejewski, Astron. J. **563**, L81 (2001).

15. K. Goździewski, E. Bois, A. J. Maciejewski, and L. Kiseleva-Eggleton, Astron. Astrophys. **378**, 569 (2001).
16. E. Hairer and G. Wanner, *Numerical Solution of a System of First Order Ordinary Differential Equations* $y' = f(x, y)$. *Fortran Odex Subroutine*, http://www.unige.ch/math (1995).
17. V. V. Kozlov, *Symmetries, Topology, and Resonances in Hamiltonian Mechanics* (Springer-Verlag, Berlin, 1996).
18. J. Laskar, Nature **338**, 237 (1989).
19. J. Laskar, Dynam. Celest. Mechan. Dynam. Astron. **56**, 191 (1993).
20. I. G. Malkin, *Theory of Motion Stability* (Nauka, Moscow, 1966).
21. A. M. Molchanov, *On the Resonant Structure of the Solar System* (Nauka, Moscow, 1973).
22. H. Poincaré, *Les Métodes nouvelles de la Mécanique céleste* (Nauka, Moscow, 1971).
23. W. H. Press, S. A. Teukolsky, W. T. Vetterling, *et al.*, *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. Press, New York, 1992).
24. G. J. Sussman and J. Wisdom, Science **241**, 433 (1988).
25. M. Tabor, *Chaos and Integrability in Nonlinear Dynamics* (Wiley, New York, 1989).
26. D. V. Treshev, Prikl. Mat. Mekh. **56**, 587 (1992).
27. N. Voglis, G. Contopoulos, and C. Efthymiopoulos, Celest. Mechan. Dynam. Astron. **73**, 211 (1999).

Translated by V. Astakhov