Astronomy Letters, Vol. 28, No. 7, 2002, pp. 423–427. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 483–488.

Do the Fundamental Constants Vary in the Course of Cosmological Evolution?

A. V. Ivanchik[1]*, E. Rodriguez[2], P. Petitjean[2,3], and D. A. Varshalovich[1]

[1] *Ioffe Physicotechnical Institute, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia*

[2] *Institut d'Astrophysique de Paris—CNRS, Paris, France*

[3] *LERMA—Observatoire de Paris, France*

Received January 30, 2002

Abstract—The possible cosmological variation of the proton-to-electron mass ratio $\mu = m_p/m_e$ was estimated by measuring the H_2 wavelengths in the spectra of distant quasars. We analyze high-resolution ($FWHM \approx 7$ km s^{-1}) spectra of the two damped Lyman-α systems at redshifts $z_{abs} = 2.3377$ and 3.0249 observed in the spectra of the quasars Q 1232+082 and Q 0347−382, respectively. Our analysis yielded the most conservative estimate for the possible variation of μ in the past ~10 Gyr, $\Delta\mu/\mu = (5.7 \pm 3.8) \times 10^{-5}$. Since the significance of this result does not exceed 1.5σ, further observations are needed to increase the statistical significance. This is the most stringent limit on the possible cosmological variation of μ to date.

Key words: *theoretical and observational cosmology, quasar spectra, fundamental constants*

INTRODUCTION

Current theories of fundamental interactions (SUSY GUT, Superstrings/M-theory, and others) predict variations of the fundamental physical constants as the Universe evolves. First, the theories predict variations of the coupling constants with increasing particle interaction energy (the so-called running constants). This effect has been reliably confirmed in high-energy accelerator experiments. Thus, for example, the fine-structure constant $\alpha = e^2/\hbar c$ is 1/137.036 at low energies and 1/128.896 at energy ~90 GeV (Vysotsky *et al.* 1996). This effect must be taken into account when considering the very early Universe.

Another prediction of the current theories is that the low-energy limits of the physical constants can vary in the course of cosmological evolution and take on different values at different points in space–time. There are several reasons for such variations. Thus, in multidimensional theories (Kaluza–Klein models, p-brane models, and others), the variations in fundamental physical constants are a direct result of the cosmological evolution of extra-dimensional subspace. In several theories (e.g., Superstring), the variations in constants result from the evolution of the vacuum state (a vacuum condensate of some scalar field or quintessence).

Experimental detection of such variations in constants would be a great step forward in understanding and faithfully describing Nature. The publication by Webb *et al.* (2001) on the detection of a possible variation in the fine-structure constant, $\Delta\alpha/\alpha = (-0.72 \pm 0.18) \times 10^{-5}$, at an epoch corresponding to redshifts $0.5 < z < 3.5$ caused quite a stir. The method used by these authors is based on a simultaneous measurement of the variations in the wavelengths of a large number of transitions for various species, which significantly reduced statistical errors. However, it was more difficult to estimate the systematic errors than in the previously used method where the fine structure of lines of each species was measured separately (see, e.g., Ivanchik *et al.* 1999). In any case, this intriguing result must be checked independently by using a different method.

THE PROTON-TO-ELECTRON MASS RATIO

Here, we check a possible cosmological variation $\Delta\mu/\mu$, where μ is the proton-to-electron mass ratio at the epoch $z = 2.3$–3.0. Note that a variation in α basically suggests a variation in μ, because any kind of interaction inherent in a given particle contributes to its observed mass. This implies that any variations in interaction parameters must cause a variation in the particle mass and, consequently, in μ. Unfortunately, the physical mechanism responsible for the generation of the proton and electron masses

*E-mail: iav@astro.ioffe.rssi.ru

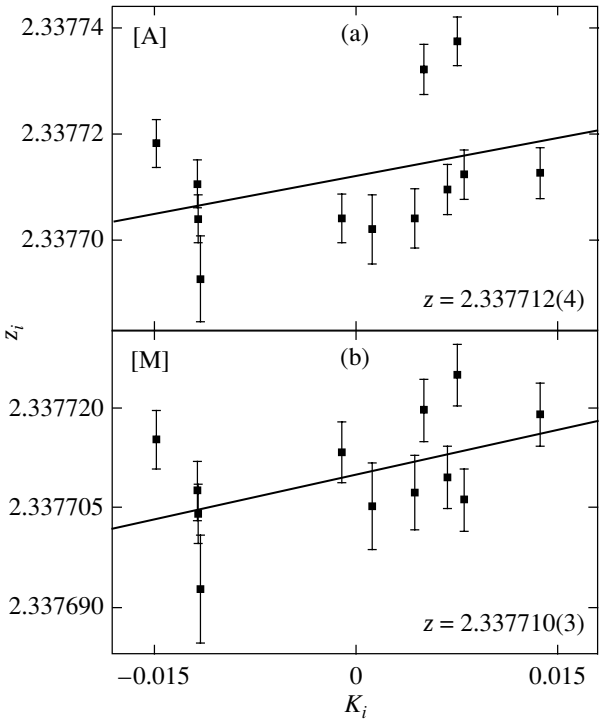

Fig. 1. The results of z_i–K_i regression analysis for the H_2 system at $z_{abs} = 2.3377$ in the Q 1232+082 spectrum. The laboratory wavelengths were taken from (a) Abgrall *et al.* (1993) and (b) Morton and Dinerstein (1976).

is still unclear. Therefore, the exact functional dependence $\mu(\alpha)$ is unknown. Nevertheless, there are several models that allow the electromagnetic contribution to the proton and electron masses to be estimated (see, e.g., Gasser and Leutwyler (1982) or Damour and Polyakov (1994), in which m_p, m_e, and α depend on the amplitude of a scalar field varying in the course of cosmological evolution). Model relations between cosmological variations of α and m_p were also derived (Calmet and Fritzsch 2001). Note that the numerical value of $\mu = m_p/m_e$ is approximately equal to the ratio of the strong interaction constant $g^2/(\hbar c) \approx 14$ to the electromagnetic interaction constant $\alpha = e^2/\hbar c \approx 1/137$, where g is the effective coupling constant calculated from the pion–nucleon scattering amplitude at low energies.

The current proton-to-electron mass ratio has been measured with a relative accuracy of 2×10^{-9}; $\mu = 1836.1526670(39)$ (Mohr and Tailor 2000). Laboratory metrological measurements rule out any significant variations in constants on a short time scale but do not rule out their variations on a cosmological time scale ($\sim 10^{10}$ yrs). Moreover, one cannot exclude the possibility of a difference between constants in spatially separated regions of the Universe. This can be verified only through astrophysical observations of extragalactic objects. By measuring the wavelengths of absorption lines in the spectra of high-redshift quasars, we can directly estimate the possible deviation of fundamental physical constants (in particular, of μ and α) at the epoch when the absorption spectra were formed, i.e., $\sim$10–13 Gyr ago.

The most stringent estimate of the possible cosmological variation in μ to date was obtained by Potekhin *et al.* (1998), $\Delta\mu/\mu = (-10 \pm 8) \times 10^{-5}$.

SENSITIVITY COEFFICIENTS

The method used here to determine the possible cosmological variation of μ was proposed by Varshalovich and Levshakov (1993). It is based on the fact that the wavelengths of electron-vibro-rotational lines depend on the reduced mass of the molecule, with this dependence being different for different transitions. This makes it possible to distinguish the cosmological redshift of a line from the shift caused by a possible variation in μ. The variation in wavelength λ_i with μ can be described (for $\Delta\mu/\mu \ll 1$) by the sensitivity coefficient K_i defined as

$$K_i = \frac{\mu}{\lambda_i}\frac{d\lambda_i}{d\mu}. \quad (1)$$

The sensitivity coefficients were calculated for the Lyman and Werner bands of molecular hydrogen by Varshalovich and Levshakov (1993) and Varshalovich and Potekhin (1995).

Thus, the measured wavelength λ_i of a line formed in an absorption system at redshift z_{abs} is given by

$$\lambda_i = \lambda_i^0(1 + z_{abs})(1 + K_i\Delta\mu/\mu), \quad (2)$$

where λ_i^0 is the laboratory (vacuum) transition wavelength. This expression can be represented in terms of the redshift $z_i = \lambda_i/\lambda_i^0 - 1$ as

$$z_i = z_{abs} + bK_i, \quad (3)$$

where $b = (1 + z_{abs})\Delta\mu/\mu$.

In reality, z_i is measured with an error determined by the statistical errors of astronomical measurements of λ_i and by the errors of laboratory measurements of λ_i^0. Therefore, relation (3) is only approximate. Nevertheless, if $\Delta\mu/\mu$ is nonzero, there must be a correlation between z_i and K_i. Thus, a linear regression analysis of these quantities yields z_{abs} and b and, consequently, $\Delta\mu/\mu$ and its statistical significance.

Table 1. Parameters of H_2 lines ($z_{abs} = 2.3377$) in the Q 1232+082 spectrum

Transition	λ_i^0 [M], Å	λ_i^0 [A], Å	λ_i, Å	$\sigma(\lambda_i)$, Å	K_i
L 0–0 P(3)	1115.896	1115.895	3724.543	0.005	−0.01479
L 0–0 R(3)	1112.584	1112.583	3713.480	0.005	−0.01178
L 0–0 P(2)	1112.495	1112.459	3713.179	0.005	−0.01170
L 1–0 P(4)	1104.084	1104.084	3685.093	0.009	−0.01154
L 2–0 P(3)	1084.559	1084.562	3619.947	0.005	−0.00098
L 3–0 P(4)	1074.313	1074.314	3585.740	0.007	0.00122
L 3–0 R(4)	1070.898	1070.899	3574.344	0.006	0.00439
L 3–0 P(3)	1070.142	1070.138	3571.834	0.005	0.00511
L 3–0 R(3)	1067.478	1067.474	3562.948	0.005	0.00758
L 3–0 P(2)	1066.901	1066.899	3561.002	0.005	0.00812
L 4–0 P(4)	1060.580	1060.580	3539.908	0.005	0.00685
L 4–0 P(2)	1053.281	1053.283	3515.556	0.005	0.01369

Table 2. Parameters of H_2 lines ($z_{abs} = 3.0249$) in the Q 0347–382 spectrum

Transition	λ_i^0 [M], Å	λ_i^0 [A], Å	λ_i, Å	$\sigma(\lambda_i)$, Å	K_i
L 2–0 R(1)	1077.698	1077.697	4337.614	0.010	0.00535
L 3–0 R(2)	1064.995	1064.993	4286.483	0.015	0.00989
L 3–0 P(1)	1064.606	1064.606	4284.924	0.006	0.01026
L 3–0 R(1)	1063.460	1063.460	4280.313	0.010	0.01132
L 4–0 R(3)	1053.976	1053.977	4242.144	0.010	0.01304
L 4–0 R(2)	1051.497	1051.498	4232.175	0.020	0.01536
L 6–0 R(3)	1028.986	1028.983	4141.571	0.015	0.02262
L 7–0 R(1)	1013.434	1013.436	4078.977	0.007	0.03062
W 0–0 Q(2)	1010.941	1010.938	4068.911	0.010	−0.00686
W 0–0 R(2)	1009.030	1009.023	4061.215	0.015	−0.00503
L 9–0 R(1)	992.022	992.013	3992.754	0.010	0.03796
W 1–0 Q(2)	987.978	987.974	3976.492	0.010	0.00394
W 1–0 R(1)	985.651	985.636	3967.087	0.007	0.00626
L 10–0 P(1)	982.834	982.834	3955.814	0.010	0.04053
L 12–0 R(3)	967.674	967.675	3894.798	0.008	0.04386
W 2–0 Q(2)	967.278	967.279	3893.194	0.010	0.01301
W 2–0 Q(1)	966.097	966.094	3888.423	0.007	0.01423
W 3–0 Q(1)	947.425	947.422	3813.255	0.008	0.02176

OBSERVATIONS AND RESULTS

High-resolution (FWHM ≈ 7 km s^{-1}) quasar spectra obtained with the 8.2-m VLT/UVES KUEYEN telescope (ESO) were used to check the possible variation in μ. We analyzed two H_2 absorption systems at $z_{abs} = 2.3377$ in the spectrum of Q 1232+082 (Petitjean *et al.* 2000) and at $z_{abs} = 3.0249$ in the spectrum of Q 0347–382 (UVES commissioning data; see D'Odorico *et al.* 2001).

The H_2 Absorption System at $z = 2.3377$ in the Spectrum of Q 1232+082

More than 50 lines of molecular hydrogen (with a signal-to-noise ratio from 10 to 14) can be identified in the wavelength range 3400–3800 Å. For

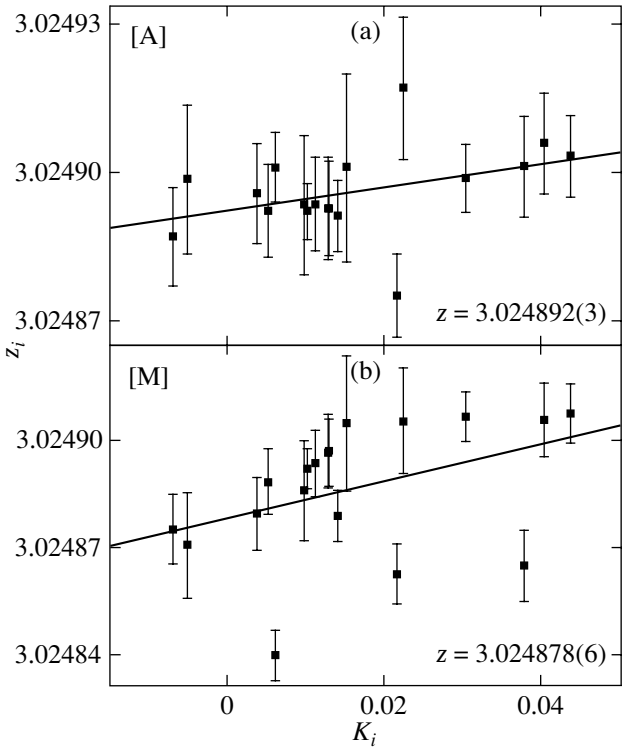

Fig. 2. Same as Fig. 1 for the H_2 system at $z_{\rm abs} = 3.0249$ in the Q 0347−382 spectrum.

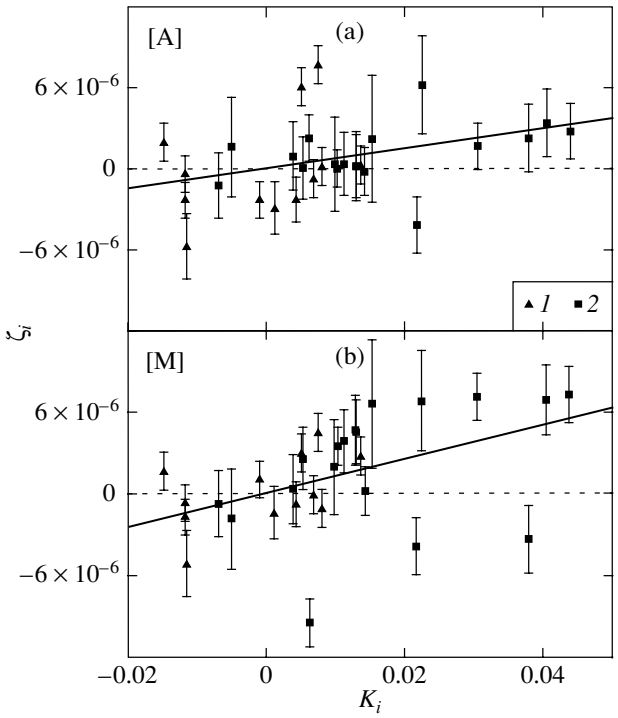

Fig. 3. The results of simultaneous $\zeta_i - K_i$ regression analysis for the systems at $z_{\rm abs} = 2.3377$ (*1*) and 3.0249 (*2*). The laboratory wavelengths were taken from (a) Abgrall *et al.* (1993) and (b) Morton and Dinerstein (1976).

our analysis, we carefully selected lines that satisfied the following conditions: (i) isolated, (ii) unsaturated, and (iii) unblended. In this system, only 12 lines satisfy these conditions; their parameters are given in Table 1. The observed wavelengths λ_i were measured with an average accuracy of ∼5 mÅby taking into account the number of points in the line profile, the spectral resolution, and the signal-to-noise ratio (Eq. (A14) from Bohlin *et al.* (1983)). For the laboratory wavelengths, we used two independent data sets: λ_i^0 [M] (Morton and Dinerstein 1976) and λ_i^0 [A] (Abgrall *et al.* 1993; see also Roncin and Launay 1994). Figure 1 shows the results of our linear regression analysis of z_i as a function of K_i for these two sets of laboratory wavelengths. They correspond to $\Delta\mu/\mu = (14.4 \pm 11.4) \times 10^{-5}$ [A], and $\Delta\mu/\mu = (13.2 \pm 7.4) \times 10^{-5}$ [M].

The H_2 Absorption System at $z = 3.0249$ in the Spectrum of Q 0347−382

This H_2 system was first detected and investigated by Levshakov *et al.* (2002). More than 80 lines of molecular hydrogen (with a signal-to-noise ratio from 10 to 40) can be identified in the wavelength range 3600−4600 Å. We independently reanalyzed this spectrum. For our analysis, we selected 18 H_2 lines that satisfied the above conditions. The observed wavelengths λ_i were measured with an average accuracy of ∼10 mÅ. Parameters of these lines are given in Table 2. Figure 2 shows the results of our linear regression analysis of z_i as a function of K_i for these lines. They correspond to $\Delta\mu/\mu = (5.8 \pm 3.4) \times 10^{-5}$ [A] and $\Delta\mu/\mu = (12.2 \pm 7.3) \times 10^{-5}$ [M]. It should be noted that three points in Fig. 2b deviate from the regression line by more than 3σ. Two of them corresponding to the L 9−0 R(1) and W 1−0 R(1) transitions were marked by Morton and Dinerstein (1976) as a blended line and as a line with a weak continuum. The third point corresponding to the W 3−0 Q(1) transition deviates in Figs. 2a and 2b, which may be a result of an undetectable blend in the quasar spectrum. We do not reject these points, because all of them satisfy the above conditions for line selection from quasar spectra.

Simultaneous Analysis

A simultaneous analysis of the H_2 lines from the two systems allows us to increase the statistical significance both through an increase in the total number of lines involved in our statistical analysis and

through a broadening of the range of sensitivity coefficients. Figure 3 shows the results of our linear regression analysis of ζ_i as a function of K_i for all 30 lines from the two systems. Here, ζ_i is the reduced line redshift:

$$\zeta_i = \frac{z_i - \overline{z}}{1 + \overline{z}}, \quad (4)$$

where $\overline{z}$ is $z_{\rm abs}$ of the absorption system under consideration that corresponds to the set of laboratory wavelengths used.

Our simultaneous analysis yielded the following estimates (for the two sets of laboratory wavelengths):

$$\Delta\mu/\mu = (5.7 \pm 3.8) \times 10^{-5}\ [\text{A}],$$
$$\Delta\mu/\mu = (12.5 \pm 4.5) \times 10^{-5}\ [\text{M}].$$

The measurement errors of the laboratory wavelengths are $\sim$1.5 mÅ. Therefore, the systematic error due to uncertainties in laboratory wavelength measurements is of the order of 2×10^{-5}, in agreement with the error in $\Delta\mu/\mu$ obtained from the two independent sets of laboratory wavelengths.

CONCLUSIONS

Our results may be considered as a hint of a possible cosmological variation in μ. Further measurements are needed to reach a more definitive conclusion. Nevertheless, we obtained the most stringent estimate to date of a possible cosmological variation in μ between zero redshift and $z = 2$–3.

Measurements of more H_2 absorption systems at high redshifts are required to improve the result. The spectra of the quasars PKS 0528−250, Q 0347−382, and Q 1232+082 are most suitable for this purpose. Observations of these quasars with a high resolution (FWHM $\sim$ 7 km s^{-1}) and a high signal-to-noise ratio ($>$30) would make the conclusion more definitive.

In addition, it is desirable to increase the accuracy of the H_2 laboratory wavelengths, because the contribution of their statistical errors is comparable to the statistical errors of astronomical observations.

ACKNOWLEDGMENTS

The observations were carried out with the 8.2-m KUEYEN telescope (UVES/VLT) operated by the European Southern Observatory (Parana, Chili). We wish to thank C. Ledoux for primary reduction of the spectra and A. Potekhin for helpful discussions. A. Ivanchik and D. Varshalovich are grateful to the Russian Foundation for Basic Research, project nos. 02-02-06096mas and 02-02-16278a, for support. A. Ivanchik is grateful for the opportunity to visit the IAP CNRS.

REFERENCES

1. H. Abgrall, E. Roueff, F. Launay, *et al.*, Astron. Astrophys., Suppl. Ser. **101**, 273 (1993).
2. R. C. Bohlin, J. K. Hill, E. B. Jenkins, *et al.*, Astrophys. J., Suppl. Ser. **51**, 277 (1983).
3. X. Calmet and H. Fritzsch, hep-ph/0112110 (2001).
4. T. Damour and A. M. Polyakov, Nucl. Phys. B **423**, 532 (1994).
5. S. D'Odorico, M. Dessauges-Zavadsky, and P. Molaro, Astron. Astrophys. **368**, L21 (2001).
6. J. Gasser and H. Leutwyler, Phys. Rep. **87**, 77 (1982).
7. A. V. Ivanchik, A. Y. Potekhin, and D. A. Varshalovich, Astron. Astrophys. **343**, 439 (1999).
8. S. A. Levshakov, M. Dessauges-Zavadsky, S. D'Odorico, and P. Molaro, Astrophys. J. **565**, 696 (2002).
9. P. J. Mohr and B. N. Taylor, Rev. Mod. Phys. **72** (2), 351 (2000).
10. D. C. Morton and H. L. Dinerstein, Astrophys. J. **204**, 1 (1976).
11. P. Petitjean, R. Srianand, and C. Ledoux, astro-ph/0011437 (2000).
12. A. Y. Potekhin, A. V. Ivanchik, D. A. Varshalovich, *et al.*, Astrophys. J. **505**, 523 (1998).
13. J.-Y. Ronchin and F. Launay, J. Phys. Chem. Ref. Data, No. 4 (1994).
14. D. A. Varshalovich and C. A. Levshakov, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 231 (1993) [JETP Lett. **58**, 237 (1993)].
15. D. A. Varshalovich and A. Y. Potekhin, Space Sci. Rev. **74**, 259 (1995).
16. M. I. Vysotskiĭ, V. A. Novikov, L. B. Okun', and A. N. Rozanov, Usp. Fiz. Nauk **166** (5), 539 (1996) [Phys. Usp. **39**, 503 (1996)].
17. J. K. Webb, M. T. Murphy, V. V. Flambaum, *et al.*, Phys. Rev. Lett. **87**, 091301 (2001).

Translated by A. Ivanchik

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 428–433. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 489–494.

A Simple Model for the Evaporation of Black Holes at Final Stages

S. O. Alexeyev[1,2], A. Barrow[1], G. Bowdole[1], M. V. Sazhin[2], and O. S. Khovanskaya[2*]

[1] *Institut de Physique Nucleair, Grenoble, France*

[2] *Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119899 Russia*

Received March 4, 2002

Abstract—We present a simple model for the evaporation of primordial black holes at final stages with the formation of a relic remnant with a mass of $1–10^3\ m_{\rm Pl}$. The model takes into account the conservation of energy and the impossibility of passing through the state with the minimum possible mass. These relic remnants may account for a substantial fraction of dark matter in the Universe. © *2002 MAIK "Nauka/Interperiodica".*

Key words: *black holes, evaporation at final stages*

INTRODUCTION

Presently, one of the most interesting and puzzling questions in modern theoretical physics concerns the final stage of Hawking evaporation of primordial black holes (PBHs). According to the standard scenario and the Hawking formula (Novikov and Frolov 1986), they must be evaporated completely. At the same time, several models (Polnarev and Khlopov 1981; Khlopov *et al.* 1985; MacGibbon 1987; MacGibbon and Carr 1991; Manko and Markov 1993; Alexeyev and Pomazanov 1997; Alexeyev and Sazhin 1998; Alexeyev *et al.* 2001) predict a lower limit on the possible black-hole mass. Here, we deal with simple modifications of the black-hole evaporation law when the minimum possible black-hole mass is constrained. This fact may play a major role in cosmology, solving the problem of dark-matter candidates. It should be noted that various ways of stopping the evaporation in Lovelock gravity were examined by Myers and Simon (1988) and, using string-like series in curvature and possible cosmological implications of this fact, by Barrow *et al.* (1992). The relationship of the PBH spectrum to parameters of the early Universe was considered previously (Alexeyev *et al.* 2002). Our subsequent papers are devoted to more realistic models of PBH evaporation.

THE GROUND STATE OF A QUASI-CLASSICAL MODEL

Since the general theory of relativity is not renormalizable, it cannot be directly quantized. Additional approaches must be used to obtain quantum gravity. One of these approaches involves using M theory, which gives 11-dimensional supergravity in the low-energy limit and the five known string theories in 10-dimensional space, as the basic model and the subsequent compactification into our 4-dimensional space–time. In the perturbative approach, the inferred effective gravity (which gives general relativity in a zero approximation) is a series in curvature with an additional scalar field (dilaton) of the form

$$S = \frac{1}{16\pi}\int d^4x\sqrt{-g} \qquad (1)$$
$$\times\left[m_{\rm Pl}^2\left(-R + 2\partial_\mu\phi\partial^\mu\phi\right) + \lambda e^{-2\phi}S_{\rm GB} + \ldots\right],$$

where R is the scalar curvature, $\phi = \phi(r)$ is the dilaton, $m_{\rm Pl}$ is the Planck mass, and λ is the string coupling constant. In addition to the Einstein term, the action includes the dilatonic field and high-order curvature corrections. The two-loop correction given by

$$S_{\rm GB} = R_{ijkl}R^{ijkl} - 4R_{ij}R^{ij} + R^2$$

is called the Gauss–Bonnet term.

Such actions may be considered as an intermediate step between general relativity and quantum gravity, which currently does not exist in finished form. However, a number of interesting results can be obtained even in terms of this model. One of them is a Gauss–Bonnet black hole[1] (Mignemi and Stewart 1993; Kanti *et al.* 1996; Torii *et al.* 1997). The stability of such objects at all singular points against small temporal perturbations was shown by

*E-mail: khovansk@xray.sai.msu.ru

[1] It should be noted that, since all the results being discussed have been obtained recently, as yet there is no standard terminology.

Khovanskaya (2002). In addition, according to the theory of catastrophes, the Gauss–Bonnet dilatonic black holes are stable against strong perturbations (Torii and Maeda 1998). Therefore, the solution found may be assumed to be stable everywhere. An important characteristic of a Gauss–Bonnet black hole is the existence of a minimum possible black-hole size (or mass). There is no static, asymptotically flat, spherically symmetric black hole with a mass smaller than the minimum mass $M_{\min}$ that corresponds to the horizon radius

$$r_h = \sqrt{\lambda}\sqrt{4\sqrt{6}}\,\phi_h(\phi_\infty),$$

where $\phi_h(\phi_\infty)$ is the dilaton value on the horizon, which depends on the dilaton value at infinity (an additional external model parameter). This state is called the ground state for a quasi-classical model. We consider a quasi-classical state that becomes the ground state when the model is quantized.

The question of whether the transition from the next to last (first excited) state to the ground state is possible in principle must be clarified.

We consider a diagonal quasi-Schwarzschild metric of the form[2]

$$ds^2 = \Delta(r)dt^2 - \frac{\sigma^2(r)}{\Delta(r)}dr^2 - r^2(d\theta^2 + \sin^2\theta d\psi^2). \quad (2)$$

At position $r_{h\min}$ (see Fig. 1), the asymptotic form of metric (2) is

$$\Delta = \mathrm{const}_1 \times \sqrt{r - r_{h\min}},$$
$$\sigma = \mathrm{const}_2 \times \sqrt{r - r_{h\min}}.$$

Consequently, $R_{ijkl}R^{ijkl} \approx \mathrm{const}_3 \times (r - r_{h\min})^{-6}$; i.e., the curvature invariant diverges, and this is a nonintegrable singularity. At the same time, the asymptotics on the horizon is regular in all the remaining states and shows a quasi-Schwarzschild behavior:

$$\Delta = d_1(r - r_h) + d_2(r - r_h)^2 + \ldots, \quad (3)$$
$$\sigma = s_0 + s_1(r - r_h) + \ldots,$$

where $(r - r_h) \ll 1$, s_0 and r_h are free parameters.

The probability of the transition from the first excited state to the ground state with a minimum mass is (Branoff and Brill 1999)

$$P = \mathrm{const} \times e^{S_{r_h} - S_{r_{h\min}}} = \mathrm{const} \times e^{-S_{r_{h\min}}}$$
$$\propto \mathrm{const} \times e^{-\frac{1}{(r - r_{h\min})^5}} = \mathrm{const} \times e^{-\infty} = 0.$$

Thus, this transition is forbidden and the black hole will never reach the ground state in the course of its evaporation.

[2]In the metric and in our subsequent calculations, we use a system of units with $\hbar = c = G = 1$.

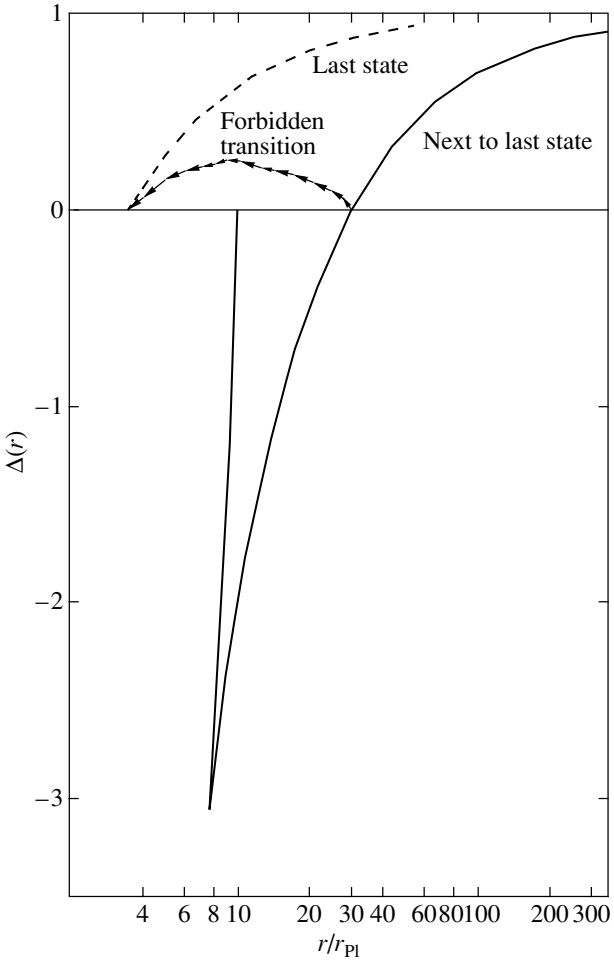

Fig. 1. An illustration of the last transition: the next to last state is characterized by the existence of a regular horizon with a finite quasi-Schwarzschild asymptotics; since the last state (a minimum black hole) is a singularity, the transition to this state is forbidden by quantum mechanics.

THE DECELERATION OF EVAPORATION

To stop the evaporation before reaching the ground state $M_{\min}$ requires including the condition for stopping the evaporation in the classical emission model.[3] Since the mass of the relic remnant of a black hole becomes comparable to the mass of its emitted particles, the condition for stopping the evaporation can be obtained from the fact that the black hole cannot emit more matter than it contains.

In the standard approach, the mass of the emitted particles is much smaller than M of the emitting object itself. The smaller is the black-hole mass, the more intense is the emission. The evaporation rate is

[3]The classical emission model assumes the mass of the emitted particles to be much smaller than the mass of the emitting object itself.

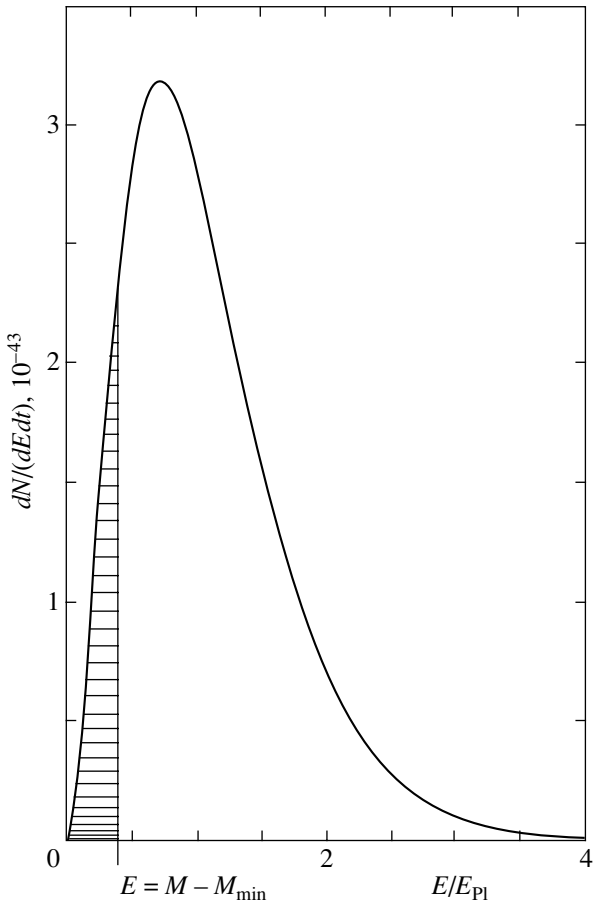

Fig. 2. A simple model for the deceleration and stopping of black-hole evaporation at final stages involves introducing a cutoff of the Planck spectrum by the law of energy conservation. A particle of the corresponding energy can be emitted if the particle energy lies to the left of the cutoff line $E = M - M_{\min}$ (hatched region) and cannot be emitted if it lies to the right, because its energy exceeds the total energy of the system. In this case, the black hole can pass into the forbidden region. Thus, although the black hole temperature increases, the energy of the emitted particles decreases until the evaporation is completely stopped.

inversely proportional to M^2:

$$-\frac{dM}{dt} \approx 4 \times 10^{-5} \left(\frac{m_{\rm Pl}}{M}\right)^2 \frac{m_{\rm Pl}}{t_{\rm Pl}} f, \qquad (4)$$

where $f = 1.023h(1/2) + 0.420h(1) + 0.048h(2)$ is a function of the number of degrees of freedom for particles with spins 1/2, 1, and 2, respectively (Novikov and Frolov 1986).

Let us consider the emission process (4) by taking into account the fact that the mass of the emitted particle for a black hole with a mass of the order of the minimum mass becomes comparable to that of the emitting object itself. In this case, it would be natural to impose the condition that during its evaporation, the black hole cannot emit more matter than it contains. This condition follows from the existence of a forbidden transition between the ground state of the black hole and its excited states. In the standard model of Hawking emission, the mass-loss rate of the black hole increases to infinity as its mass tends to zero (4).[4] If conservation of energy is taken into account, the mass of the black hole decreases as it evaporates (in the simplest way) and its temperature increases, but the mean energy of the emitted particle and its frequency decrease. Thus, the classical Planck spectrum is cut off by the condition that the energy E of the emitted particle should not exceed $M - M_{\min}$. When emitting particles, the system successively passes to new states without reaching the forbidden ground state $E = M - M_{\min} = 0$ (Fig. 2). This condition can be taken into account by inserting the Heaviside function (H) in the formula for emission. Thus, we obtain

$$\frac{d^2N}{dEdt} = \frac{\Gamma_s(M \times E)}{2\pi} \frac{H(M - M_{\min} - E)}{e^{8\pi ME} - (-1)^{2s}}, \qquad (5)$$

where $\Gamma_s(M \times E)$ is the Starobinsky–Page function (Starobinsky 1973; Starobinsky and Churilov 1973; Page 1976a, 1976b). This function depends on the mass, energy, and spin of the corresponding particle emitted by the black hole as follows:

$$\Gamma_{s\rm boson} = \left[\frac{(l-s)!(l+s)!}{(2l)!(2l+1)!!}\right]^2 \times \prod_{n=1}^{l}\left[1 + \frac{16}{n^2}(ME)^2\right] 8(ME)\left[2ME\right]^{2l+1},$$

$$\Gamma_{s\rm fermion} = \left[\frac{(l-s)!(l+s)!}{(2l)!(2l+1)!!}\right]^2 \times \prod_{n=1}^{l+1/2}\left[1 + \frac{64}{(2n-1)^2}(ME)^2\right]\left[2ME\right]^{2l+1},$$

where l and s are the quantum numbers and $(M \times E) \ll 1$. Below, we take into account the contribution of the $l = s$ mode alone, which is dominant.

Given expression (5), the emission law (4) is modified as follows:

$$-\frac{dM}{dt} = \frac{1}{2\pi}\int_0^{M-M_{\min}} dE \frac{\Gamma_s(M \times E)E}{e^{8\pi ME} - (-1)^{2s}}, \qquad (6)$$

where the integration is no longer over all particle energies but only up to a finite limit.

Let us consider the emission law separately for particles of different spins at $M - M_{\min} \ll 1$. At final

[4] This is the reason why the evaporation is important precisely for low-mass black holes.

evaporation stages, the corresponding rates of decrease in the black-hole mass can be represented as the joining of analytic asymptotics and series with the corresponding domains of convergence (Fig. 3).

For particles of zero spin,

$$-\frac{dM}{dt} = \frac{1}{512\pi^5}\frac{1}{M^2} \times \int_0^{8\pi M(M-M_{\min})} dx\frac{x^3}{e^x-1} = \frac{1}{\pi^2}M(M-M_{\min})^3\left(\frac{1}{3} - \pi M(M-M_{\min}) + \sum_{k=1}^{\infty}\frac{B_{2k}(8\pi M(M-M_{\min}))^{2k}}{(2k+3)(2k)!}\right). \quad (7)$$

For particles of spin 1/2,

$$-\frac{dM}{dt} = \frac{1}{8192\pi^5}\frac{1}{M^2} \times \int_0^{8\pi M(M-M_{\min})} dx\frac{x^3}{e^x+1} = \frac{1}{1024\pi^5}\frac{1}{M^2}e^{-(4\pi M(M-M_{\min})} \times \left(\sum_{k=0}^{\infty}\frac{E_{2k}(4\pi M(M-M_{\min}))^{2k+4}}{(2k+4)(2k)!} + \sum_{k=0}^{\infty}\frac{E_{2k}\gamma(2k+4,4\pi M(M-M_{\min}))}{(2k+4)(2k)!}\right). \quad (8)$$

For particles of spin 1,

$$-\frac{dM}{dt} = \frac{1}{73728\pi^7}\frac{1}{M^2} \times \int_0^{8\pi M(M-M_{\min})} dx\frac{x^5}{e^x-1} = \frac{4}{9\pi^2} \times M^3(M-M_{\min})^5\left(\frac{1}{5} - \frac{2}{3}\pi M(M-M_{\min}) + \sum_{k=1}^{\infty}\frac{B_{2k}(8\pi M(M-M_{\min}))^{2k}}{(2k+5)(2k)!}\right). \quad (9)$$

For particles of spin 2,

$$-\frac{dM}{dt} = \frac{1}{29491200\pi^9}\frac{1}{M^2} \times \int_0^{8\pi M(M-M_{\min})} dx\frac{x^9}{e^x-1} = \frac{16}{225\pi^2} \times M^5(M-M_{\min})^7\left(\frac{1}{7} - \frac{1}{2}\pi M(M-M_{\min}) + \sum_{k=1}^{\infty}\frac{B_{2k}(8\pi M(M-M_{\min}))^{2k}}{(2k+7)(2k)!}\right). \quad (10)$$

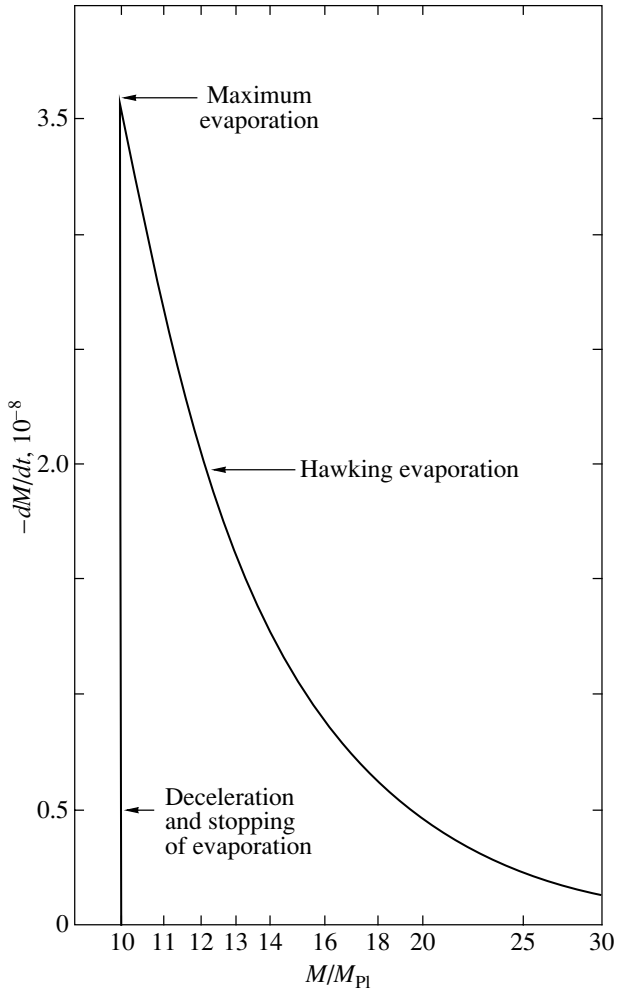

Fig. 3. The black hole evaporation law with energy conservation: the standard Hawking part where $-dM/dt \sim 1/M^2$; the stopping of evaporation at final stages upon reaching the state with a minimum mass.

The domain of convergence for the series in Eq. (8) is $|8\pi M(M-M_{\min})| < \pi$; E_{2k} are such Euler numbers that $1/\cosh(t)$ is a generating function for them:

$$\frac{1}{\cosh(t)} = \sum_{n=0}^{\infty} E_n\frac{t^n}{n!},$$

$\gamma(n,\alpha)$ is an incomplete gamma function.

In Eqs. (7), (9), and (10), the domain of convergence is $|8\pi M(M-M_{\min})| < 2\pi$ and B_{2k} are such Bernoulli numbers that $1/(e^t-1)$ is a generating function for them:

$$\frac{1}{e^t-1} = \sum_{n=0}^{\infty} B_n\frac{t^n}{n!}.$$

In the high-mass limit ($M(M-M_{\min}) \gg 1$), the evaporation rates of particles with different spins are

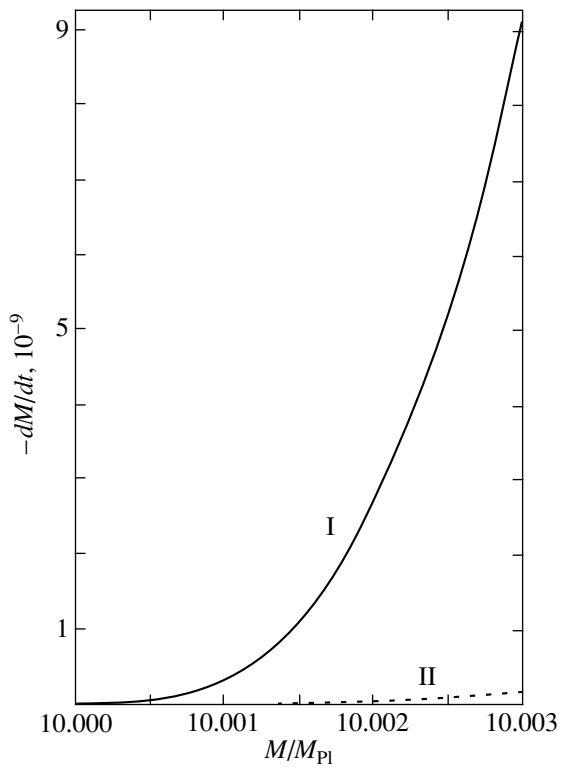

Fig. 4. The black hole evaporation law with energy conservation at final stages when the difference in the emission probabilities of particles with different spins shows up most clearly: the emission rates of bosons (I) and fermions (II).

virtually indistinguishable. Indeed, using an asymptotic representation of the integrals,

$$\int_0^{8\pi M(M-M_{\rm min})} dx\frac{x^3}{e^x\pm 1}$$
$$\approx \int_0^{\infty} dx\frac{x^3}{e^x\pm 1} - \int_{8\pi M(M-M_{\rm min})}^{\infty} dx x^3 e^{-x},$$

and taking into account the fact that the Page function in Eq. (5) is simply $\Gamma_s(M\times E) = M^2E^2$ in the high-mass limit, irrespective of the spin of the emitted particles, we obtain for bosons ($s = 0, 1, 2$)

$$-\frac{dM}{dt} = \frac{1}{128\pi^3}\frac{1}{M^2}$$
$$\times \int_0^{8\pi M(M-M_{\rm min})} dx\frac{M^2x^3}{e^x-1} \approx \frac{1}{128\pi^3}$$
$$\times\left(\frac{\pi^4}{15} - e^{-A}(A^3+3A^2+6A+6)\right)$$

and for fermions

$$-\frac{dM}{dt} = \frac{1}{128\pi^3}\frac{1}{M^2}$$
$$\times \int_0^{8\pi M(M-M_{\rm min})} dx\frac{M^2x^3}{e^x+1} \approx \frac{1}{128\pi^3}$$
$$\times\left(\frac{7\pi^4}{120} - e^{-A}(A^3+3A^2+6A+6)\right),$$

where $A = 8\pi M(M-M_{\rm min})$.

Thus, at infinity, i.e., in the classical high-mass approximation, the ratio of the boson and fermion emission rates is 8/7.

For $M(M-M_{\rm min}) \ll 1$, it would be natural to use a different approximation:

$$-\frac{dM}{dt}\bigg|_{s=0} \approx \frac{1}{3\pi^2}M(M-M_{\rm min})^3, \qquad (11)$$

$$-\frac{dM}{dt}\bigg|_{s=1/2} \approx \frac{1}{16\pi}M^2(M-M_{\rm min})^4, \qquad (12)$$

$$-\frac{dM}{dt}\bigg|_{s=1} \approx \frac{4}{45\pi^2}M^3(M-M_{\rm min})^5, \qquad (13)$$

$$-\frac{dM}{dt}\bigg|_{s=2} \approx \frac{16}{1575\pi^2}M^5(M-M_{\rm min})^7. \qquad (14)$$

The evaporation behavior at final stages for bosons and fermions is shown in Fig. 4.

Relations (11)–(14) lead us to conclude that particles with an integer spin will be emitted at final stages of black hole evaporation with a high probability.

CONCLUSIONS

We have presented a simple model to describe the evaporation of primordial black holes at final stages with the formation of a relic remnant with a mass of $1-10^3\ m_{\rm Pl}$. These relic remnants may account for a substantial fraction of dark matter in the Universe. In subsequent papers, we plan to investigate the probability of experimentally detecting these PBHs in a more realistic model.

ACKNOWLEDGMENTS

We wish to thank A.A. Starobinsky for a helpful discussion and for valuable remarks. S. Alexeyev is grateful to the Institut de Physique Nucleair (Grenoble) for creating favorable conditions for collaboration. A. Barrow and G. Bowdole are grateful to the Sternberg Astronomical Institute (Moscow State University) for a kind invitation. This study was supported in part by the Program "Russian Universities—Basic Research," grant no. 990777.

REFERENCES

1. S. O. Alexeyev and M. V. Pomazanov, Phys. Rev. D **55**, 2110 (1997).
2. S. O. Alexeyev and M. V. Sazhin, Gen. Relativ. Gravit. **8**, 1187 (1998).
3. S. O. Alexeyev, M. V. Sazhin, and M. V. Pomazanov, Int. J. Mod. Phys. D **10**, 225 (2001).
4. S. O. Alexeyev, M. V. Sazhin, and O. S. Khovanskaya, Pis'ma Astron. Zh. **28**, 163 (2002) [Astron. Lett. **28**, 139 (2002)].
5. J. D. Barrow, E. J. Copeland, and A. R. Liddle, Phys. Rev. D **46**, 645 (1992).
6. P. R. Branoff and D. R. Brill, gr-qc/9811079.
7. P. Kanti, N. E. Mavromatos, J. Rizos, *et al.*, Phys. Rev. D **54**, 5049 (1996).
8. M. Yu. Khlopov, B. A. Malomed, and Ya. B. Zeldovich, Mon. Not. R. Astron. Soc. **215**, 575 (1985).
9. O. Khovanskaya, submitted to Grav. Cosmol. (2002).
10. J. H. MacGibbon, Nature **329**, 308 (1987).
11. J. H. MacGibbon and B. Carr, Astrophys. J. **371**, 447 (1991).
12. V. I. Manko and M. A. Markov, Phys. Lett. A **172**, 331 (1993).
13. S. Mignemi and N. R. Stewart, Phys. Rev. D **47**, 5259 (1993).
14. R. C. Myers and J. Z. Simon, Phys. Rev. D **38**, 2434 (1988).
15. I. D. Novikov and V. P. Frolov, *Physics of Black Holes* (Nauka, Moscow, 1986).
16. D. N. Page, Phys. Rev. D **13**, 198 (1976a).
17. D. N. Page, Phys. Rev. D **14**, 3260 (1976b).
18. A. G. Polnarev and M. Yu. Khlopov, Astron. Zh. **58**, 706 (1981) [Sov. Astron. **25**, 406 (1981)].
19. A. A. Starobinsky, Zh. Éksp. Teor. Fiz. **64**, 48 (1973) [Sov. Phys. JETP **37**, 28 (1973)].
20. A. A. Starobinsky and S. M. Churilov, Zh. Éksp. Teor. Fiz. **65**, 3 (1973) [Sov. Phys. JETP **38**, 1 (1974)].
21. T. Torii, H. Yajima, and K. Maeda, Phys. Rev. D **55**, 739 (1997).
22. T. Torii and K. Maeda, Phys. Rev. D **58**, 084004 (1998).

Translated by V. Astakhov

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 434–437. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 495–498.

On the Equation of State for the Λ field

D. I. Podolsky*

Landau Institute for Theoretical Physics, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 117940 Russia

Received March 5, 2002

Abstract—The recently detected accelerated expansion of the Universe is related to the existence of a new type of matter called the Λ field or quintessence. Constraints were obtained on its equation of state from the absence of clustering of this matter on scales much smaller than the cosmological horizon. The question of how these constraints affect the possibility of fitting the accelerated expansion by such cosmological models as the Chaplygin gas model is discussed. © *2002 MAIK "Nauka/Interperiodica"*.

Key words: *Expansion of Universe, Λ field, equation of state*

Presently, the interpretation of the measured redshift dependence of the photometric distance for type Ia supernovae as evidence for the existence of a new type of matter in the Universe appears to have been universally accepted. From a cosmological point of view, this matter acts as the Λ term in the Einstein equations (Garnavich *et al.* 1998; Perlmutter *et al.* 1998; Riess *et al.* 1998; Schmidt *et al.* 1998). A natural question is how to reconstruct the Lagrangian of this matter from observational data.

To date, little information is available on this type of matter: (1) the cosmological effect related to the Λ term (below referred to as the Λ field[1]) is known to be isotropic; and (2) the Λ field is not clustered appreciably; i.e., it is spatially homogeneous on all scales where the distribution of nonrelativistic matter (baryons and cold dark matter) is nonuniform: $R \lesssim 30h^{-1}$ Mpc, where h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}. The first statement leads us to believe that we are dealing with the scalar mode of some quantum field theory and suggests that its effective Lagrangian is (Boisseau *et al.* 2000):

$$L = \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - F(\phi)R - V(\phi). \qquad (1)$$

It turns out (Starobinsky 1998) that if we have the redshift dependence of the photometric distance $D_L(z)$ for some class of objects, e.g., type Ia supernovae and/or the correlation functions $\langle\delta(0)\delta(z)\rangle$ of the density nonrelativistic matter component at hand, we can reconstruct the functions $F(\phi)$ and $V(\phi)$.

Recall briefly what the basic idea is by assuming that $F(\phi) = 0$ and by following Starobinsky (1998). In a system with a scalar field and a nonrelativistic matter component, the cosmology-defining equations are

$$8\pi GV(\phi) = aH\frac{dH}{da} + 3H^2 - \frac{3}{2}\Omega_{m,0}H_0^2\left(\frac{a_0}{a}\right)^3, \qquad (2)$$

$$4\pi Ga^2H^2\left(\frac{d\phi}{da}\right)^2 = -aH\frac{dH}{da} - \frac{3}{2}\Omega_{m,0}H_0^2\left(\frac{a_0}{a}\right)^3, \qquad (3)$$

where H_0 is the Hubble parameter at $z = 0$; i.e., presently, $\Omega_{m,0} = \dfrac{8\pi G\epsilon_m}{3H_0^2}$ is the density of the dust-like matter component in dimensionless units at $z = 0$, $\dfrac{d}{dt} = aH\dfrac{d}{da}$. If the expression for $H(z)$ is known, then the dependence $a(\phi)$ must be derived from Eq. (3) and substituted in Eq. (2) to obtain the potential $V(\phi)$. In turn, $H(z)$ can be derived from $D_L(z)$ or $\delta(z)$ by using expressions from Starobinsky (1998) (see also Huterer and Turner 1999; Nakamura and Chiba 1999):

$$H(z) = \left(\frac{d}{dz}\left(\frac{D_L(z)}{1+z}\right)\right)^{-1}, \qquad (4)$$

$$\frac{H^2(z)}{H^2(0)} = \frac{(1+z)^2\delta'^2(0)}{\delta'^2(z)} - 3\Omega_{m,0}\frac{(1+z)^2}{\delta^2(z)}\int_0^z\frac{\delta|\delta'|}{1+z}dz. \qquad (5)$$

Equation (5) is valid only for $F(\phi) = 0$; otherwise, the situation becomes more complicated (Boisseau *et al.* 2000).

*E-mail: podolsky@itp.ac.ru

[1] For the variable Λ term, Huey *et al.* (1999) also proposed the name "quintessence."

1063-7737/02/2807-0434$22.00

The same method can be used to reconstruct its equation of state $p_\Lambda = f(\epsilon_\Lambda)$ without concretizing the effective Lagrangian of the Λ field. The corresponding equations are

$$\frac{8\pi G}{3}\epsilon_\Lambda = H_0^2\Omega_{m,0}(1+z)^3 - H^2(z), \quad (6)$$

$$4\pi G f(\epsilon_\Lambda) = -4\pi G\epsilon_\Lambda - \frac{3}{2}\Omega_{m,0}H_0^2(1+z)^3 + H(z)(1+z)\frac{dH(z)}{dz}. \quad (7)$$

If $H(z)$ is known, then we can deduce $z(\epsilon_\Lambda)$ from Eq. (6) and derive the equation of state for the Λ field by substituting it in Eq. (7).

Actual experimental data currently correspond to $z \lesssim 1$ in these equations if $H(z)$ is determined from $D_L(z)$ in experiments with type Ia supernovae (Garnavich *et al.* 1998; Perlmutter *et al.* 1998; Riess *et al.* 1998; Schmidt *et al.* 1998). Equation (5) could be applied by using the abundance of rich galaxy clusters as a function of redshift, $n(z)$, to calculate $\delta(z)$. However, the corresponding data are not yet sufficiently representative and accurate. That is why any information that allows the class of equations of state containing the equation of state for the Λ field to be restricted will be useful. Let us show what constraints on the equation of state for the Λ field follow from the requirement that there be no field clustering on scales much smaller than the cosmological horizon.

The equations of the perturbation theory in cosmology where a two-component fluid (nonrelativistic matter and the Λ field) acts as the matter are (see, e.g., Solov'eva and Starobinsky 1985)

$$\dot\Phi + H\Phi = 4\pi G\left(\epsilon_m v_m + (\epsilon_\Lambda + f(\epsilon_\Lambda))\, v_\Lambda\right), \quad (8)$$

$$\frac{d}{dt}\left(\frac{\dot v_\Lambda}{f'(\epsilon_\Lambda)}\right) - 3\frac{d}{dt}(Hv_\Lambda) + \frac{k^2}{a^2}v_\Lambda = 3\dot\Phi + \frac{d}{dt}\left(\frac{\Phi}{f'(\epsilon_\Lambda)}\right), \quad (9)$$

$$\dot v_m = \Phi, \quad (10)$$

$$\dot\delta = \frac{d}{dt}\left(\frac{\delta\epsilon_m}{\epsilon_m}\right) = 12\pi G(\epsilon_\Lambda + f(\epsilon_\Lambda)) \times (v_\Lambda - v_m) - \frac{k^2}{a^2}v_m, \quad (11)$$

$$-3\frac{d}{dt}(Hv_m) + \frac{k^2}{a^2}v_m = 3\dot\Phi, \quad (12)$$

$$4\pi G(\delta\epsilon_m + \delta\epsilon_\Lambda) = -\frac{k^2}{a^2}\Phi, \quad (13)$$

$$\frac{\delta\epsilon_\Lambda}{\epsilon_\Lambda + f(\epsilon_\Lambda)} = \frac{1}{f'(\epsilon_\Lambda)}\frac{d}{dt}(v_\Lambda - v_m), \quad (14)$$

where v is the gauge-invariant velocity potential and δ is the Bardeen gauge-invariant energy-density perturbation (Bardeen 1980). We consider Friedmann metric perturbations in the longitudinal gauge, i.e., $ds^2 = (1+2\Phi)dt^2 - a^2(t)(1-2\Psi)\delta_{lm}dx^l dx^m$, where $l, m = 1, 2, 3$, and assume a dependence of all perturbations on the spatial coordinates of the form $\exp(ik_j x^j)$, $f'(\epsilon_\Lambda) = \beta_\Lambda^2 = \dfrac{dp_\Lambda}{d\epsilon_\Lambda}$. For this physical system, $\Phi = \Psi$.

The absence of Λ-field clustering on small scales in a broad sense implies that the equation for nonrelativistic-matter density perturbations on scales much smaller than the Hubble radius $R = H^{-1}$ is the same as it would be in the absence of the Λ field:

$$\ddot\delta + 2H\dot\delta - 4\pi G\epsilon_m\delta = 0. \quad (15)$$

It can be derived from Eqs. (8)–(14) of the perturbation theory as follows. From Eq. (11), we have $a^2\dot\delta = 12\pi G(\epsilon_\Lambda + f(\epsilon_\Lambda))(v_\Lambda - v_m)a^2 - k^2 v_m$. Differentiating this equation with respect to t and using Eqs. (10) and (14), we obtain

$$\ddot\delta + 2H\dot\delta = -\frac{k^2}{a^2}\Phi + 12\pi G(1 + f'(\epsilon_\Lambda)) \times \dot\epsilon_\Lambda(v_\Lambda - v_m) + 24\pi GH(\epsilon_\Lambda + f(\epsilon_\Lambda)) \times (v_\Lambda - v_m) - 12\pi G f'(\epsilon_\Lambda)\delta\epsilon_\Lambda. \quad (16)$$

Clearly, this equation does not always reduce to Eq. (15) in the short-wavelength limit ($\dfrac{k}{aH} \gg 1$).

We will compare the first term on the right-hand side with the remaining terms. Since $\dot\Phi \sim H\Phi$, we have $v_m \sim \dfrac{a^2}{k^2}(H\Phi)$ from Eq. (12) in the short-wavelength limit. Assume that Eq. (9) in this limit gives

$$v_\Lambda \sim \frac{a^2}{k^2}\left(3\dot\Phi + \frac{d}{dt}\left(\frac{\Phi}{f'(\epsilon_\Lambda)}\right)\right). \quad (17)$$

As can be easily seen, this is the case if

$$\left(\frac{aH}{k}\right)^2\left|\frac{1}{f'(\epsilon_\Lambda)}\right| \lesssim 1, \quad (18)$$

$$\left(\frac{aH}{k}\right)\frac{a}{k}\left|\frac{d}{dt}\left(\frac{1}{f'(\epsilon_\Lambda)}\right)\right| \lesssim 1 \quad (19)$$

(here, $x \lesssim 1$ implies that x cannot be too large compared to unity).

The satisfaction of these conditions is sufficient for the first term on the right-hand side of Eq. (16) to be

dominant. Indeed, let us compare the first term with the third term on the right in Eq. (16):

$$\pi G(\epsilon_\Lambda + f(\epsilon_\Lambda))(v_\Lambda - v_m) < \left(\frac{aH}{k}\right)^2 \times \Phi\left(\frac{d}{dt}\left(\frac{1}{f'(\epsilon_m)}\right) + \frac{H}{f'(\epsilon_m)} - 3H\right) \ll \frac{k^2}{Ha^2}\Phi. \quad (20)$$

Similarly, it can be shown that the second term on the right is small compared to the first term (the identity $\dot{\epsilon}_\Lambda = -3H(\epsilon_\Lambda + f(\epsilon_\Lambda))$ can be used for this purpose).

Let us now show that the inequality $|\delta\epsilon_\Lambda| \ll |\delta\epsilon_m|$, which essentially implies the absence of Λ-field clustering in a narrow sense, holds in the short-wavelength limit when conditions (18) and (19) are satisfied. Indeed, we have from Eq. (14)

$$|\delta\epsilon_\Lambda| = \left|\frac{(\epsilon_\Lambda + f(\epsilon_\Lambda))}{f'(\epsilon_\Lambda)}\frac{d}{dt}(v_\Lambda - v_m)\right| \sim \frac{H^2}{\pi G}\left(\frac{a}{k}H\right)^2\Phi. \quad (21)$$

If we now use Eq. (13), then we will see that the smallness of $\delta\epsilon_\Lambda$ relative to $\delta\epsilon_m$ is of the order of $\left(\frac{aH}{k}\right)^4$.

Let us try to understand how greatly constraints (18) and (19) allow the class to which the theory describing the Λ field belongs to be narrowed. First, these constraints become meaningful when a quantum field theory of type (1) begins to appear hydrodynamical on large scales. In general, for an arbitrary quantum field theory, the nondiagonal components of the stress tensor are not small compared to its diagonal components (i.e., $\sigma_{ik} \neq p\delta_{ik}$). This implies that there are no perturbations propagating at the speed of sound $\sqrt{dp/d\epsilon}$ in the system or that they are of no crucial importance. Accordingly, the hydrodynamic degrees of freedom are not physical, reconstructing the equation of state ceases to be of value, and we must return to the formulation of the inverse cosmological problem in the form given by Starobinsky (1998).

Hydrodynamics faithfully describes the actual situation if the mean free path for perturbations in the corresponding quantum field theory is small compared to the Hubble radius $1/H$. In a theory of type (1), if we take the expansion of the potential energy near the minimum

$$V(\phi) \approx V_0 + \frac{m^2\phi^2}{2} + \frac{\lambda\phi^4}{4} + \cdots, \quad (22)$$

this condition appears as

$$l_{kin} \sim \frac{1}{n\sigma} \sim \frac{m}{\Lambda}\frac{m^2}{\lambda^2} \ll \frac{1}{H}, \quad (23)$$

where n is the perturbation density, σ is the scattering cross section in the ϕ^4 theory, and Λ is the energy density of the Λ field.

This is a stringent criterion: if we believe that the accelerated dynamics of the Universe is currently provided by the same scalar field that is responsible for the inflationary stage (a simple inflationary model with a potential of form (22) whose parameters are well known: $\lambda \sim 10^{-12}$ and $m \sim 10^{13}$ GeV), then constraints (18) and (19) are meaningless. Nevertheless, the freedom is still large: (i) a more complex theory than (1), for example, a multicomponent scalar field, may have provided the inflation; (ii) or inflaton is not responsible for the present accelerated dynamics of the Universe.

Below, we assume that condition (23) is satisfied and that hydrodynamics well describes perturbations in the Λ field. As is easy to see, since $(aH/k)^2$ is small (currently available experimental data strongly suggest only the absence of Λ-field clustering on scales <30 Mpc), constraints (18) and (19) give little for ideal fluid models of the Λ field with the equation of state $p_\Lambda = \omega\epsilon_\Lambda$, where $\omega \leq -1/3$. However, such models must be essentially of a theoretical-field nature. Indeed, the speed of sound $\frac{dp_\Lambda}{d\epsilon_\Lambda} = \omega$ squared is negative for such models, implying an exponential increase in the perturbation amplitude in such hydrodynamics.

From this viewpoint, Chaplygin gas models (Kamenshchik *et al.* 2001; see also the recent paper by Bento *et al.* 2002) with the equation of state $p_\Lambda = -A/\epsilon_\Lambda^n$, where A and $n > 0$ are constants. are of great interest. The constraints on A and n that follow from conditions (18) and (19) are meaningful:

$$\left(\frac{aH}{k}\right)^2 \frac{\Lambda^{n+1}}{nA} \lesssim 1, \quad (24)$$

$$\left(\frac{aH}{k}\right)^2 \frac{n+1}{nA}\Lambda^n\left|\Lambda - \frac{A}{\Lambda^n}\right| \lesssim 1. \quad (25)$$

Assuming the energy dominance $\epsilon_\Lambda + f(\epsilon_\Lambda) \geq 0$ for the Λ field, we find from expression (25) using (24) that

$$(n+1)\left(1 - \frac{1}{n}\left(\frac{aH}{k}\right)^2\right) \lesssim 1. \quad (26)$$

Thus, the exponent n in the equation of state for Chaplygin gas cannot be too large compared to unity.

ACKNOWLEDGMENTS

I wish to thank A.A. Starobinsky for the discussions during which the problem in question emerged. This work was supported by the Russian Foundation for Basic Research (project nos. 00-15-96699 and 02-02-16817).

REFERENCES

1. J. Bardeen, Phys. Rev. D **22**, 1882 (1980).
2. M. C. Bento, D. Bertolami, and A. A. Sen, gr-qc/0202064 (2002).
3. B. Boisseau, G. Esposito-Farese, D. Polarski, and A. Starobinsky, Phys. Rev. Lett. **85**, 2236 (2000).
4. P. M. Garnavich, S. Jha, P. Challis, *et al.*, Astrophys. J. **509**, 74 (1998).
5. G. Huey, L. Wang, R. Duve, *et al.*, Phys. Rev. D **59**, 063005 (1999).
6. D. Huterer and M. S. Turner, Phys. Rev. D **60**, 81301 (1999).
7. A. Kamenshchik, U. Moschella, and V. Pasquier, Phys. Lett. B **511**, 265 (2001).
8. T. Nakamura and T. Chiba, Mon. Not. R. Astron. Soc. **306**, 696 (1999).
9. S. Perlmutter, G. Aldering, M. Della, *et al.*, Nature **391**, 51 (1998).
10. A. G. Riess, A. V. Filippenko, P. Challis, *et al.*, Astron. J. **116**, 1009 (1998).
11. B. P. Schmidt, N. B. Suntzeff, M. M. Phillips, *et al.*, Astrophys. J. **507**, 46 (1998).
12. L. V. Solov'eva and A. A. Starobinskiĭ, Astron. Zh. **62**, 625 (1985) [Sov. Astron. **29**, 367 (1985)].
13. A. A. Starobinsky, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 721 (1998) [JETP Lett. **68**, 757 (1998)].

Translated by V. Astakhov

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 438–442. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 499–504.

Magnetic Fields of Active Galactic Nuclei and Quasars: Redshift Dependence

Yu. N. Gnedin[1]* and N. A. Silant'ev[2]

[1] *Pulkovo Observatory, Russian Academy of Sciences, Pulkovskoe shosse 65, St. Petersburg, 196140 Russia*

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

Received September 10, 2001; in final form, December 27, 2001

Abstract—The observed reduction in the fraction of quasars with polarized radiation as the redshift increases is explained by an increase in the magnetic-field strength in accretion disks at high z. The emerging Faraday rotation causes the radiation to be depolarized. This mechanism allows the magnetic fields in accretion disks to be estimated at several hundred gauss. We give simple asymptotic formulas that describe the Faraday depolarization in optically thick accretion disks. © *2002 MAIK "Nauka/Interperiodica"*.

Key words: *active galactic nuclei, quasars, and radio galaxies*

INTRODUCTION

Broadband optical polarimetry can serve as a good tool for determining the magnetic-field strengths in the immediate vicinity of active galactic nuclei (AGNs) and quasars. Gnedin and Silant'ev (1978, 1997); Silant'ev (1993); and Silant'ev *et al.* (2000) proposed a new method for estimating the magnetic fields of astrophysical objects. The essence of this method is to allow for the overall effect of Faraday rotation in electron stellar atmospheres and accretion disks with magnetic fields.

Faraday rotation produces new polarimetric effects. First, overall linear polarization emerges even in a spherically symmetric scattering circumstellar envelope or in the atmosphere of a spherical star in the presence of a magnetic field. Second, a characteristic wavelength dependence of polarization arises, with the position of the maximum depending on the magnetic-field distribution in the atmosphere. Third, the radiation is depolarized in those cases where polarization appears in the absence of a magnetic field (a nonspherical electron circumstellar envelope or stellar atmosphere, an accretion disk, a hot gaseous jet, etc. (see Dolginov *et al.* 1995).

Agol and Blaes (1996) and Agol *et al.* (1998) performed detailed calculations of polarized-radiation transfer in electron-scattering accretion disks by taking into account Faraday rotation. They also allowed for the contribution of absorption to the total atmospheric opacity. Their main conclusion is that a magnetic field reduces the degree of polarization compared to accretion-disk models without magnetic fields (see Koratkar and Blaes (1999) for a review).

Information on the structure of the global magnetic field in a quasar or AGN can also be obtained from polarimetric radio observations, particularly through interferometry. Recently, the VLBI technique has yielded data on the magnetic-field strength and geometry in the hot plasma around quasars and AGNs on the parsec or even subparsec scale (Udompraset *et al.* 1997; Ishwara-Chandra *et al.* 1998; Saikia and Kulkarni 1998; Pentericci *et al.* 2000). Our prime objective is to use the method of allowance for overall Faraday rotation that we developed to determine the magnetic-field strength in the immediate vicinity of an AGN or a quasar. It turns out that a redshift (z) dependence of the AGN and quasar magnetic fields being determined can be established in principle from available polarimetric observations.

REDSHIFT DEPENDENCE OF THE NUMBER OF QUASARS WITH POLARIZED RADIATION

The observations of several authors (Impey *et al.* 1991; Wills *et al.* 1992; Carilli *et al.* 2000) show that the number of extragalactic sources (AGNs, quasars, huge infrared galaxies) with linearly polarized radiation significantly decreases with increasing redshift z. Impey *et al.* (1991) found that, depending on the redshift, the fraction of quasars with polarized radiation is distributed as follows:

$$0.43 \pm 0.09 \quad (0 < z < 1), \qquad (1)$$
$$0.36 \pm 0.10 \quad (1 < z < 2),$$

*E-mail: gnedin@gao.spb.ru

1063-7737/02/2807-0438$22.00

$$0.22 \pm 0.16 \quad (2 < z < 3).$$

In turn, Wills *et al.* (1992) established that the fraction of quasars with high polarization, $p > 3\%$, has the distribution

$$0.59 \pm 0.06 \quad (z < 1), \qquad 0.24 \pm 0.06 \quad (z > 1). \tag{2}$$

Thus, it follows from distribution (2) that the number of quasars with strongly polarized radiation at high ($z > 1$) redshifts decreases by more than a factor of 2. Finally, Carilli *et al.* (2000) give the following distribution of strongly polarized extragalactic sources:

$$\begin{aligned} &\geq 10\% \quad (0 < z < 1.5), \\ &\geq 2\% \quad (2 < z < 2.5). \end{aligned} \tag{3}$$

Such a reduction in the number of quasars and galactic nuclei with high polarization appears to depend on the selection effect. However, the authors above believe that this effect cannot completely account for the observed reduction in the number of objects.

Two factors act against the selection effect. First, the number of quasars itself increases with redshift, reaching a maximum at $z \approx 3$. For example, having analyzed several surveys, Maday (1999) established that the density of bright quasars in a comoving coordinate system greatly increases near $z \approx 3$. This factor also causes the number of objects with substantial polarization of their radiation to increase. Second, one might expect an increase in polarization at high ($z > 1$) redshifts. The point is that Thomson scattering by electrons is the most efficient generation mechanism of polarized radiation. For high redshifts, a shorter wavelength corresponds to the observed optical wavelength λ in the frame of a quasar at rest, $\lambda_0 = \lambda/(1+z)$. The radiation spectrum shifts to the ultraviolet and even to the X-ray range at very high z. Since the absorption of waves with shorter wavelengths in plasma is weaker, one might expect the radiation from accretion disks around galactic nuclei at high redshifts to be more polarized than that from nuclei at $z \approx 1$.

Using the standard frequency dependence of the free–free absorption cross section σ, we can obtain the following redshift dependence of the expected polarization:

$$\begin{aligned} p &\propto \frac{\sigma_T}{\sigma_{ff}} \propto (1+z)^3/\lambda_0^3, \qquad h\nu \gg kT_\mathrm{e}, \\ p &\propto \frac{\sigma_T}{\sigma_{ff}} \propto (1+z)^2/\lambda_0, \qquad h\nu \ll kT_\mathrm{e}, \end{aligned} \tag{4}$$

where σ_T is the Thomson scattering cross section.

Both these effects, which can cause the fraction of quasars with polarized radiation to increase, at least compensate for much of the selection effect.

Below, we use data from Wills *et al.* (1992) and Carilli *et al.* (2000) and assume that the redshift dependence of the number of highly polarized quasars is given by the relation

$$\frac{N(z \leq 1)}{N(z > 1)} \approx 0.2\text{–}0.4. \tag{5}$$

THE STANDARD THEORY FOR THE GENERATION OF POLARIZED RADIATION IN THE IMMEDIATE VICINITY OF AN AGN OR A QUASAR

One of the fundamental components in the model of a quasar or an AGN is currently believed to be an optically thick accretion disk, which is considered as the source of quasar radiation. The continuum radiation from such a disk is linearly polarized due to Thomson scattering by electrons. The disk polarization depends on the disk inclination to the line of sight. When a geometrically thin disk is seen almost edge-on, the polarization of a disk with an electron atmosphere reaches 11.7% (the well-known Sobolev–Chandrasekhar effect) and the electric-vector oscillation direction lies in the disk plane.

When light is scattered in an optically thin, disk-like shell, the direction of the electric vector is parallel to the shell minor axis, while the polarization is proportional to the shell optical depth and can reach several percent.

The theory of polarized-radiation transfer in an electron atmosphere developed by Chandrasekhar (1950) and Sobolev (1960) underlies the standard model of polarization generation. The intensity of the radiation $I(\mu) = I_l(\mu) + I_r(\mu)$ emerging from an optically thick, plane-parallel atmosphere is the sum of the intensities of radiations polarized parallel (l) and perpendicular (r) to the plane containing the normal to the disk and the line of sight:

$$\begin{aligned} I_l(\mu) &= \frac{3}{8\pi}\frac{F}{\sqrt{2}} q H_l(\mu), \\ I_r(\mu) &= \frac{3}{8\pi}\frac{F}{\sqrt{2}} H_r(\mu)(\mu + c), \end{aligned} \tag{6}$$

where $\mu = \cos\vartheta$, ϑ is the angle between the line of sight and the normal to the disk, $q = 0/68980$, $c = 0.87294$, and F is the emergent flux,

$$F = 2\pi \int_0^1 d\mu \mu I(\mu). \tag{7}$$

The functions $H_l(\mu)$ and $H_r(\mu)$ are the solutions to the standard Ambartsumian–Chandrasekhar nonlinear equation

$$H(\mu) = 1 + \mu H(\mu) \int_0^1 d\eta \frac{\psi(\eta)}{\mu + \eta} H(\eta) \qquad (8)$$

with $\psi_l(\eta) = 3(1 - \eta^2)/4$ and $\psi_r(\eta) = 3(1 - \eta^2)/8$ substituted for $\psi(\eta)$. At $\mu = 0$, $H_l(0) = H_r(0) = 1$. If $\mu = 1$, then $H_l(1) = 3.4695$ and $H_r(1) = 1.27797$.

Since the problem is symmetric, only the Stokes parameter $Q(\mu) = I_l(\mu) - I_r(\mu)$ is nonzero and the polarization is

$$p(\mu) = \frac{Q(\mu)}{I(\mu)} = \frac{qH_l(\mu) - H_r(\mu)(\mu + c)}{qH_l(\mu) + H_r(\mu)(\mu + c)}. \qquad (9)$$

For an edge-on ($\mu = 0$) disk, the polarization is $p = (q - c)/(q + c) = -0.1171$. The minus sign implies that the electric-field oscillations are perpendicular to the plane containing the line of sight and the normal to the atmospheric surface.

THE POLARIZATION OF RADIATION FROM AN ACCRETION DISK WITH A MAGNETIC FIELD: ALLOWANCE FOR FARADAY ROTATION

Previously (Gnedin and Silant'ev 1978, 1997; Silant'ev 1994), we generalized the Sobolev–Chandrasekhar theory to an electron atmosphere with a magnetic field. We showed how Faraday rotation changes the spectral and angular distributions of the polarized radiation emerging from a magnetized, plane-parallel electron atmosphere. One of the main effects is a sharp decrease in polarization due to Faraday depolarization. Detailed calculations of polarized-radiation transfer for a magnetic field perpendicular to the disk plane were performed by several authors (Silant'ev 1994; Agol and Blaes 1996; Gnedin and Silant'ev 1997; Agol *et al.* 1998; Koratkar and Blaes 1999).

The Faraday rotation angle at Thomson optical depth τ can be written as (see Gnedin and Silant'ev 1997)

$$\psi_F = \frac{1}{2}\delta\tau\cos\theta, \qquad (10)$$

$$\delta = \frac{3}{4\pi}\frac{\lambda}{r_e}\frac{\omega_B}{\omega} \cong 0.8\lambda^2\mu(\mathrm{m})B(G),$$

where $\omega = 2\pi\nu$ is the radiation cyclic frequency, $\omega_B = eB/m_e c$ is the cyclotron frequency, $r_e = e^2/m_e c^2 \cong 2.82 \times 10^{-13}$ cm, and θ is the angle between the line of sight and the magnetic field.

Although obtaining a precise pattern for the radiation emerging at different angles to the normal to the disk surface requires numerical methods of calculation, the asymptotic formulas for the Stokes parameters when $\delta \gg 1$ can be represented in a simple analytic form (see Silant'ev 1994):

$$I(\mu) = \frac{F}{2\pi H_1} H(\mu), \qquad (11)$$

$$Q(\mu) = -\frac{F}{2\pi H_1}\frac{b}{2}\cdot\frac{1 - \mu^2}{1 + \delta^2\cos^2\theta}, \qquad (12)$$

$$U(\mu) = -\frac{F}{2\pi H_1}\frac{b}{2}\cdot\frac{(1 - \mu^2)\delta\cos\theta}{1 + \delta^2\cos^2\theta}. \qquad (13)$$

Here, $b = 0.18274$, the function $H(\mu)$ satisfies Eq. (8) with $\psi(\mu) = 3(3 - \mu^2)/16$ and corresponds to the Rayleigh phase function, and $H_1 = 1.19400$ is the first moment of $H(\mu)$ (see Chandrasekhar 1950). Note that the Stokes parameters (12) and (13) are given in a coordinate system with the x axis in the plane containing the line of sight and the normal to the disk.

For an axisymmetric magnetic-field distribution in the disk plane ($B_z = 0$, a diamagnetic disk), integration over the azimuthal angle can be easily performed to give the analytic formula

$$p(\mu, \delta) = \frac{\sqrt{Q^2 + U^2}}{I} \qquad (14)$$

$$= \frac{b}{2}\frac{1 - \mu^2}{H(\mu)}\frac{1}{\sqrt{1 + \delta^2(1 - \mu^2)}}.$$

Below, we use Eqs. (11)–(14) to estimate the Faraday depolarization as the chief cause of the reduction in the number of polarized quasars with increasing redshift z.

FARADAY DEPOLARIZATION AS A POSSIBLE CAUSE OF THE REDUCTION IN THE NUMBER OF QUASARS WITH POLARIZED RADIATION AT HIGH REDSHIFTS

Formula (5) describes the observed deficit of polarized quasars at high redshifts. What causes this effect? A plausible explanation is that this deficit results from the depolarizing action of Faraday rotation in accretion disks at high redshifts.

We assume that the magnetic field in the disk at $z < 1$ decreases through dissipation (see, e.g., Bisnovatyi-Kogan 1999) to an extent that there is virtually no Faraday rotation and that the quasar polarization is attributable to the ordinary Thomson scattering of light in a nonmagnetized hot electron atmosphere. In this case, polarization can be observed over a wide range of accretion-disk inclinations, except the directions along the normals to the disks. In contrast, the initial magnetic field of the disk

for $z > 1$, which has not yet dissipated, causes the emerging radiation to be depolarized due to Faraday rotation for all angles, except a narrow range of solid angles almost perpendicular to the magnetic field. It follows from the above theoretical formulas that for estimates, this narrow range of solid angles may be taken to be a factor of $\sim\delta$ smaller than the corresponding range of solid angles in the absence of Faraday rotation. Thus, we may set the ratio (5) equal to the ratio of the characteristic solid angles in the presence and absence of Faraday rotation:

$$\frac{N(z>1)}{N(z\le 1)} \approx 0.2\text{–}0.4 \approx \frac{\Delta\Omega_{\rm Faraday}}{\Delta\Omega} \approx \frac{\pi}{2\delta} = \frac{\pi}{1.6\lambda_0^2(\mu m)B(G)}. \quad (15)$$

Here, we use the depolarization parameter δ that corresponds to the radiation wavelength $\lambda_0 = \lambda/(1+z)$ in the quasar rest frame. As a result, we obtain an estimate of the magnetic field in the accretion disk:

$$B(G) \approx \frac{\pi(5-2.5)}{1.6\lambda^2(\mu m)}(1+z)^2. \quad (16)$$

Note that formula (16) gives only an estimate of the magnetic-field strength for an object at a given redshift z. Under no circumstances should it be considered as the establishment of an actual z dependence of the accretion-disk magnetic field.

Assuming that $\lambda \cong 0.55\ \mu m$ (which corresponds to observational data), we obtain the following magnetic-field estimate: $B \approx (30\text{–}15)(1+z)^2 G$. For a mean characteristic value of $z \ge 2$, we find the magnetic field in quasar accretion disks to be $\ge$300–150 G.

Because of the importance of these kinds of estimates, we obtain an independent estimate by setting the ratio of the mean observed polarizations for quasars at $z > 1$ and $z < 1$ equal to the theoretical ratio calculated using the above formulas. In this case, the accretion-disk planes are assumed to be randomly oriented relative to the line of sight. Using extensive observational polarization data from Impey *et al.* (1991), we find that the mean polarization of quasars is $\sim$2.7% at $z < 1$ and $\sim$2.4% at $z > 1$, i.e., it is slightly lower. When calculating these mean values, we did not use several objects whose polarization was higher than 11.7% and which could not be described in terms of the theory presented above. The theoretical polarization of the accretion disk averaged over all orientations is 3.13% when only Thomson scattering without Faraday rotation is taken into account. Recall that for objects at $z < 1$, we assume the possible magnetic field to be weak and the Faraday rotation to be of no importance. Clearly, our separation of sources into two groups with ($z > 1$) and without ($z < 1$) magnetic fields is somewhat rough. Therefore, it comes as no surprise that the mean polarization 2.7% is slightly lower than its theoretical value 3.13%. In addition, a modest contribution of the radiation emerging from the disk edge also reduces the polarization. However, these values are still close to each other, which confirms the validity of our assumption of random disk orientation and of our separation of objects into magnetic and nonmagnetic ones.

In the presence of Faraday rotation, the theoretical polarization averaged over the disk orientations generally depends on the magnetic-field distribution in the disks. Thus, for disks with magnetic fields directed along the normal to the disk, its value is $\langle p\rangle \approx b\ln(2\delta)/2\delta$. For an axisymmetric field distribution in the disk plane, $\langle p\rangle \approx 0.22b/\delta$. Recall that $b = 0.1827$. The ratio of the observed mean polarizations for the groups with $z > 1$ and $z < 1$ is 0.9. Setting this value equal to the theoretical values $\langle p\rangle/0.0313$ of this ratio, we obtain $\delta \approx 10$ and 1.4 for the vertical and horizontal (in the disk plane) magnetic fields, respectively. The estimate (16) yields $\delta \cong 8\text{–}4$.

Thus, the two independent estimates of δ (i.e., the magnetic field) lead to identical results. We have a magnetic-field estimate $B \ge 40(1+z)^2$ G for $\delta \approx 10$, which gives $B \ge 400$ G at $z > 2$. It should be noted that the above values of $\langle p\langle$ were obtained only when averaging over the random accretion-disk orientations. Clearly, they depend on z. Unfortunately, the available observational data do not allow $\langle p\rangle$ to be calculated for each z. The $\langle p\rangle = 2.4\%$ used is a z-averaged value. Since the number of objects at high z rapidly increases, this value of $\langle p\rangle$ is mainly determined by objects with high z. Therefore, the above magnetic-field estimates should be referred to the upper limit of z used in averaging the z interval.

Note that data on the rotation measures of high-redshift galaxies (Pentericci *et al.* 2000) yield an independent estimate of the accretion-disk magnetic field in the region where polarized optical radiation is generated.

The expression for the rotation measure is

$$RM = 0.8N_{\rm e}LB_z \text{ rad m}^{-2}, \quad (17)$$

where the electron number density $N_{\rm e}$ is measured in cm^{-3}, the line-of-sight magnetic field B_z is in μG, and the radiation path length L is in pc.

The typical values of these parameters are $N_{\rm e} \sim 1$ cm^{-3}, $B_z \sim 10\ \mu$G, and $L \sim 100$ pc. Assuming the magnetic flux to be conserved and given that the typical size of the region where optical radiation is generated in the accretion disk is $\sim 10^3 r_{\rm g}$, where $r_{\rm g}$ is the gravitational radius of a supermassive black hole,

we obtain the following estimate of the accretion-disk magnetic field in the generation region of optical radiation:

$$B \sim 2 \times 10^3 \text{ G} \quad (RM/10^3)(10^8 M_\odot/M_{\rm BH})^2. \quad (18)$$

The field strength (18) is close to the magnetic-field strengths derived from our independent estimates.

CONCLUSIONS

The observed reduction in the fraction of quasars with polarized radiation as the redshift increases can be explained by an increase in the magnetic-field strength of accretion disks at cosmological ($z \geq 2$) distances. We considered the case where radiation is polarized through its Thomson scattering in a plane-parallel atmosphere. The question arises as to whether such a consideration is applicable to those accretion-disk regions where optical radiation is generated. It is well known that, for example, free–free absorption may prove to be a dominant opacity mechanism in the outer regions of accretion disks. For quasars and AGNs at high ($z > 1$) redshifts, the radiation in the intrinsic frame of reference corresponds to the ultraviolet; i.e., it is generated in the accretion-disk regions closest to a supermassive black hole. In addition, allowance for absorption can occasionally lead to an increase in polarization (see the monographs by Dolginov *et al.* (1995) and Gnedin and Silant'ev (1997), as well as Agol and Blaes (1996)). However, our comparatively rough estimates show that the observed reduction in the fraction of polarized quasars with increasing z can actually be due to the stronger magnetic fields of accretion disks at large cosmological distances. Note that the z dependence of the rotation measure revealed by radio polarimetry for high-redshift galaxies also exhibits its increase with redshift (Pentericci *et al.* 2000). Generalizing our theory to accretion disks with true photon absorption involves no fundamental difficulties (Dolginov *et al.* 1995; Silant'ev 1994).

Note that our conclusion about an increase in the magnetic-field strength for $z \geq 1.5$ is also in good agreement with the results of Willott *et al.* (2000). According to these authors, many massive galaxies with active nuclei passed through the evolutionary phases during which a supermassive black hole and stellar bulge were formed. According to this study, such a phase occurs precisely at large cosmological distances, $z \sim 2$–5.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 99-02-16366, as well as the "Integration" Program, project K032, and the Federal Program "Astronomy", project no. 1.4.2.1.

REFERENCES

1. E. Agol and O. Blaes, Mon. Not. R. Astron. Soc. **282**, 965 (1996).
2. E. Agol, O. Blaes, and C. Ionescu-Zanetti, Mon. Not. R. Astron. Soc. **293**, 1 (1998).
3. G. S. Bisnovatyi-Kogan, Izv. Odess. Astron. Obs. **12**, 169 (1999).
4. C. L. Carlii, G. Miley, H. J. A. Rottgering, *et al.*, astro-ph/0008380 (2000).
5. S. Chandrasekhar, *Radiative Transfer* (Clarendon, Oxford, 1950; Inostrannaya Literatura, 1953).
6. A. Z. Dolginov, Yu. N. Gnedin, and N. A. Silant'ev, *Propagation and Polarization of Radiation in Cosmic Media* (Gordon and Breach, New York, 1995), p. 127.
7. Yu. N. Gnedin and N. A. Silant'ev, Astron. Zh. **55**, 564 (1978) [Sov. Astron. **22**, 325 (1978)].
8. Yu. Gnedin and N. A. Silant'ev, *Basic Mechanism of Light Polarization in Cosmic Media* (Hartwood Academic Publ., Amsterdam, 1997), p. 30.
9. C. D. Impey, C. R. Lawrence, and S. Tapia, Astrophys. J. **375**, 46 (1991).
10. C. H. Ishwara-Chandra, D. J. Saikia, V. K. Kapahi, and P. J. McCarthy, Preprint No. 6198 (National Center for Radio Astrophysics, 1998).
11. A. Koratkar and O. Blaes, Publ. Astron. Soc. Pac. **111**, 1 (1999).
12. P. Maday, Preprint No. 1375 (Space Telescope Science Institute, 1999).
13. L. Pentericci, W. van Rreeven, C. L. Carilli, *et al.*, astro-ph/0005524 (2000).
14. D. J. Saikia and A. P. Kulkarni, Preprint No. 7 (National Center for Radio Astrophysics, 1998).
15. N. A. Silant'ev, Astrophys. J. **419**, 294 (1993).
16. N. A. Silant'ev, J. Quant. Spectrosc. Radiat. Transf. **52**, 207 (1994).
17. N. A. Silant'ev, Yu. N. Gnedin, and T. Sh. Krymski, Astron. Astrophys. **357**, 1151 (2000).
18. V. V. Sobolev, *A Treatise on Radiative Transfer* (Van Norstrand, Princeton, 1960).
19. P. S. Udompraset, G. B. Taylor, and T. J. Roberts, Astrophys. J. Lett. **483**, L9 (1997).
20. C. J. Willott, St. Rawlings, and K. M. Blundell, in *Proceedings of Workshop on "QSO Hosts and Their Invironments," IAA, Granada, 2001,* Ed. by I. Marques; astro-ph/0104118.
21. B. J. Wills, D. Wills, M. Breger, *et al.*, Astrophys. J. **398**, 454 (1992).

Translated by V. Astakhov

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 443–453. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 505–414.

A Spectroscopic Study of the Peculiar Galaxy UGC 5600

L. V. Shalyapina[1]*, A. V. Moiseev[2], and V. A. Yakovleva[1]

[1] *Astronomical Institute, St. Petersburg State University, Bibliotechnaya pl. 2, Petrodvorets, 198904 Russia*

[2] *Special Astrophysical Observatory, Russian Academy of Sciences, Nizhnii Arkhyz, Karachai-Cherkessian Republic, 357147 Russia*

Received February 8, 2002

Abstract—We present our observations of the galaxy UGS 5600 with a long-slit spectrograph (UAGS) and a multipupil field spectrograph (MPFS) attached to the 6-m Special Astrophysical Observatory telescope. Radial-velocity fields of the stellar and gaseous components were constructed for the central region and inner ring of the galaxy. We proved the existence of two nearly orthogonal kinematic subsystems and conclude that UGC 5600 is a galaxy with an inner polar ring. In the circumnuclear region, we detected noncircular stellar motions and suspected the existence of a minibar. The emission lines are shown to originate in H II regions. We estimated the metallicity from the intensity ratio of the [N II]λ6583 and Hα lines to be nearly solar, which rules out the possibility that the polar ring was produced by the accretion of gas from a dwarf companion. © *2002 MAIK "Nauka/Interperiodica"*.

Key words: *polar-ring galaxies, UGC 5600, kinematics*

INTRODUCTION

Polar-ring galaxies (PRGs) are objects with two kinematic subsystems rotating in roughly orthogonal planes. Based on its structural features (see Fig. 1), Whitmore *et al.* (1990) listed the peculiar galaxy UGC 5600 among the most probable PRG candidates: its amorphous main body is surrounded by a broad outer ring, and the brightenings attributed to the inner ring are observed in the E–W direction on both sides of the center at a distance of $\sim 10''$. UGC 5600 is a member of a double system (VV 330). Its companion, the galaxy UGC 5609, is at $1'.4$ (about 15 kpc in projection onto the plane of the sky) to the southeast and has a similar radial velocity. Presently, two other galaxies with similar redshifts have been detected near VV 330 (Galletta *et al.* 1997); all of them may represent a group of galaxies.

Among other galaxies from the catalog of Whitmore *et al.* (1990), UGC 5600 was observed in the radio range at a wavelength of 21 cm (Richter *et al.* 1994). The $M_{\rm HI}/L_B$ ratio was found to be 0.86, characteristic of late-type galaxies. CO-line observations revealed molecular hydrogen in all structures of the galaxy (Galletta *et al.* 1997).

The spectra of the galaxy along its major and minor axes were obtained by Reshetnikov and Combes (1994). The radial-velocity curves are complex in shape. The authors suggested counterrotation in the central part of the galaxy ($|r| < 5''$) and gas rotation around its major axis. In addition, they pointed out that the galaxy is rich in gas and that the Hα emission extends to $30''$ (6 kpc).

A detailed photometric study of UGC 5600 (Karataeva *et al.* 2001) shows that this is most likely a late-type spiral (Scd) galaxy with an inner polar ring which is projected onto the galaxy main body on the northern side and is seen through it on the southern side. The structure that was taken in the catalog of Whitmore *et al.* (1990) as an outer ring represents two tightly wound spiral arms.

The final conclusion as to whether UGC 5600 belongs to PRGs can be reached only after proving the existence of two nearly orthogonal kinematic subsystems.

The distance to the galaxy is 37.6 Mpc ($H_0 = 75$ km s^{-1}Mpc^{-1}), and the scale is 0.18 kpc in $1''$.

OBSERVATIONS AND DATA REDUCTION

All of the spectroscopic data were obtained at the prime focus ($F/4$) of the 6-m Special Astrophysical Observatory (SAO) telescope. A log of observations is given in the table.

The observations with the UAGS long-slit spectrograph (Afanas'ev *et al.* 1995) were carried out in January 2000 at two slit positions: along the galaxy major axis (PA = 0°–2°) and along the polar-ring major axis (PA = 85°); according to Karataeva

*E-mail: lshal@astro.spbu.ru

1063-7737/02/2807-0443$22.00

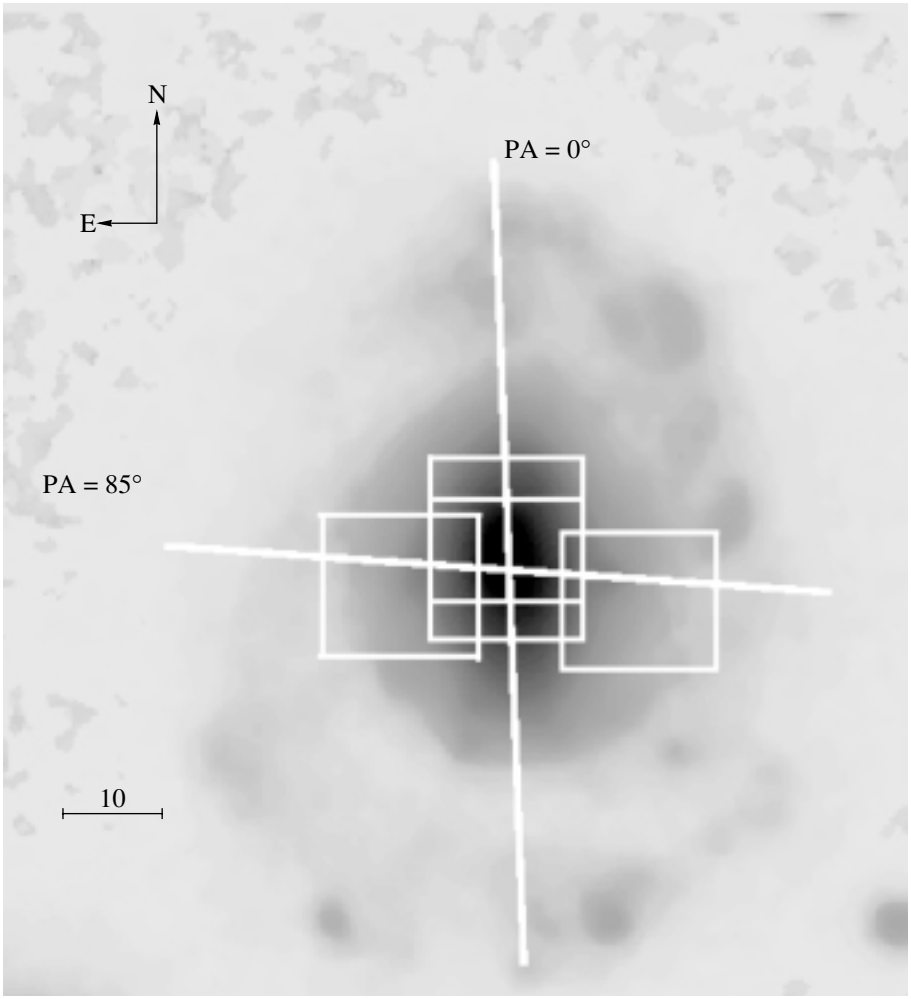

Fig. 1. A B-band image of the galaxy UGC 5600; the straight lines indicate the UAGS slit positions, and the rectangles indicate the MPFS fields.

et al. (2001), this axis passes 2″ south of the galaxy photometric center (see Fig. 1). The spectral range covered included the Hα, [N II] $\lambda\lambda$6548, 6583, and [S II] $\lambda\lambda$6716, 6730 emission lines. The detector was a Photometrics 1024 × 1024-pixel CCD array. In the observations, the spectrograph slit size was 2″ × 140″, and the angular scale along the slit was 0.″4 per pixel.

The UAGS spectra were reduced by using standard procedures from the ESO-MIDAS package. After the primary reduction, we carried out a smoothing along the slit with a 0.″8 window for the central region and a 2″ window starting from a distance of 15″ from the center. The radial velocities were measured from the centroid positions of the Gaussians fitted in the emission lines. The accuracy of these measurements was estimated from the night-sky [O I] λ6300 line to be ±10 km s^{-1}. We also measured the relative intensities and FWHMs of the above emission lines. The observed FWHMs were corrected for the instrumental profile width using the standard relation $(\mathrm{FWHM})^2 = (\mathrm{FWHM})^2_{\mathrm{obs}} - (\mathrm{FWHM})^2_{\mathrm{instr}}$.

To study the kinematics of the ionized gas and stars in the inner regions of UGC 5600 in detail, we observed the galaxy by the method of field spectroscopy with a multipupil fiber spectrograph (MPFS) (Afanas'ev *et al.* 2001) attached to the 6 m telescope. The spectrograph simultaneously takes spectra from 240 spatial elements (constructed in the form of square lenses) that form an array of 16 × 15 elements in the plane of the sky. The angular size of a single element is 1″. A description of the spectrograph is given on the Internet on the SAO page.[1] Simultaneously with galaxy spectra, we took a night-sky spectrum from an area located 4.′5 from the center of the field of view. The detector was a TK1024

[1] http://www.sao.ru/ğafan/devices/mpfs/mpfs_main.htm

A log of observations of UGC 5600

Date, instrument	Exposure time, s	Field	Seeing	Spectral range, Å	PA, field
Jan. 28, 2000	1800	$2'' \times 140''$	$2''.0$	6200–7000	85°
UAGS	1800	2×140	2.0	6200–7000	0°
	1200	2×140	2.0	6200–7000	2°
Jan. 28, 2001	1800	16×15	2.0	4300–5600	Center
MPFS	1200	16×15	2.0	4300–5600	E side
	1200	16×15	2.0	4300–5600	W side
	900	16×15	1.5	5550–6900	Center
	1200	16×15	2.0	5550–6900	E side
	1200	16×15	2.0	5550–6900	W side
Apr. 28, 2001	600	16×15	2.0	4700–5900	Center
MPFS	1200	16×15	2.0	4700–5900	+5″ to north
Aug. 11, 2001	1200	16×15	1.0	4900–6200	Center
MPFS	900	16×15	1.0	6000–7300	Center

1024×1024-pixel CCD array. The spectrograph reciprocal dispersion was 1.35 Å per pixel and the spectral resolution was ~3.5 Å. The observations were performed sequentially in two spectral ranges. The "green" range included emission lines (Hβ, [O III] $\lambda\lambda$4959, 5007) and absorption lines of the galaxy stellar population (Mg I λ5175, Fe I λ5229, Fe I+Ca I λ5270, etc.). The "red" range contained the Hα, [N II] $\lambda\lambda$6548, 6583, [S II] $\lambda\lambda$6716, 6730 emission lines.

We reduced the observations using the software developed at the SAO and running in the IDL environment. The primary reduction included debiasing, flat fielding, cosmic-ray hit removal, extraction of individual spectra from CCD images, and their wavelength calibration using the spectrum of a calibration lamp. Subsequently, we subtracted the night-sky spectrum from the galaxy spectra. The spectra of spectrophotometric standard stars were used to convert fluxes into absolute energies.

We constructed two-dimensional intensity and radial-velocity (velocity fields) maps in the Hα, Hβ, [O III] λ5007, and [N II] λ6583 lines; the emission-line profiles were also fitted with Gaussians. The accuracy of the absolute radial-velocity determination estimated from sky lines ranges from 10 to 15 km s^{-1}. The radial-velocity fields for the stellar component were constructed by the cross correlation technique (Tonry and Davis 1979) modified to work with field spectroscopy and detailed by Moiseev (2001). We used the spectral range 5200 to 5500 Å containing high-contrast lines of the galaxy stellar population. The spectra of G8–K3 III stars and the twilight sky observed on the same nights as the galaxy were taken as the radial-velocity standards. The accuracy of the radial-velocity determination is ~10 km s^{-1}.

The January 2001 observations were carried out at three different positions of the spectrograph field of view (Fig. 1). The resulting fields of velocities and emission-line intensities were combined to give a $40'' \times 16''$ total field of view. We measured the radial velocities of the stellar component only for the central $16'' \times 15''$ field, because the contrast of the stellar lines in the outermost parts decreases sharply.

In April 2001, we managed to construct a more extended radial-velocity field by using two MPFS fields (one coincided with the galaxy photometric center and the other was displaced by 5″ to the north). The resulting field of view was $16'' \times 20''$. When studying this velocity field in detail, we suspected that the central region of the galaxy, $r \sim 2''$ in size, was kinematically decoupled (see below). To check this feature, we repeated our observations of the stellar kinematics in the central region of UGC 5600 in August 2001, at ~1″ seeing. The derived $16'' \times 15''$ velocity field with a higher angular resolution was also used in our analysis. Here, all radial velocities were reduced to the solar center (heliocentric velocities).

RESULTS OF OBSERVATIONS WITH THE LONG-SLIT SPECTROGRAPH

Data on the UAGS observations are given in the first three rows of the table. In our spectra of the

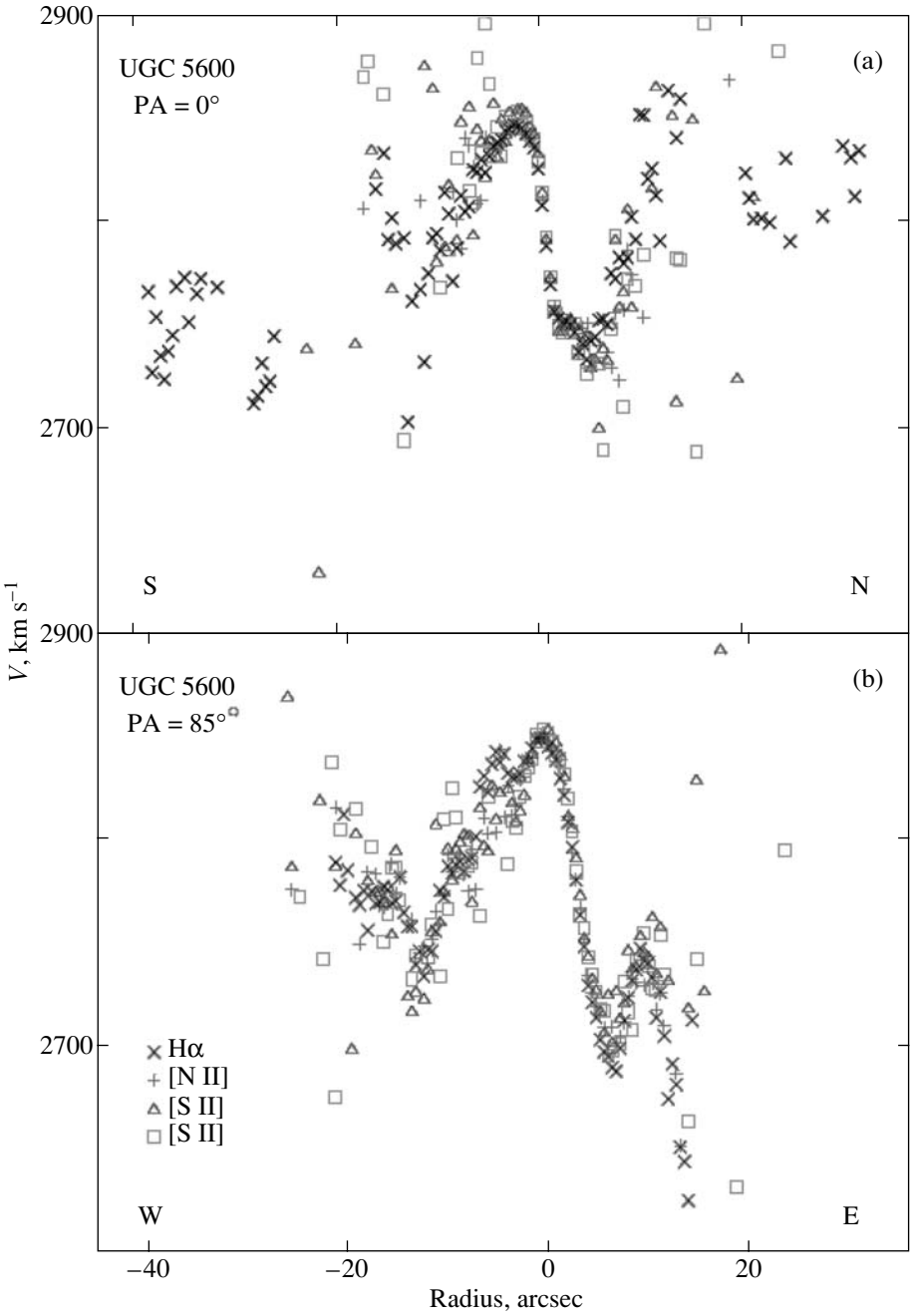

Fig. 2. Radial-velocity curves (a) along the galaxy major axis and (b) along the ring major axis.

galaxy UGC 5600 along its major axis, the Hα emission line is traceable up to distances of $\sim$7–8 kpc from the center.

The radial-velocity curves along the galaxy major axis (PA = 0°) and along the major axis of the inner ring (PA = 85°) are shown in Fig. 2a and 2b. We see from these figures that the radial velocities measured from different emission lines are equal, within the error limits. Our radial-velocity curves are similar to those in Reshetnikov and Combes (1994) and the small deviations are most likely due to differences in the spectrograph slit positions. The mean heliocentric velocity of the photometric center is 2795 ± 3 km s^{-1}, which is lower than its value in Reshetnikov and Combes (1994) by 28 km s^{-1}.

The radial-velocity curve along the galaxy major axis is complex in shape. The curve exhibits a small rectilinear segment where the velocity increases from 0 to 55 km s^{-1}; then, at a distance of 3″–4″ from the center, the scatter of points increases and further out, the mean radial velocity decreases. Reshetnikov and

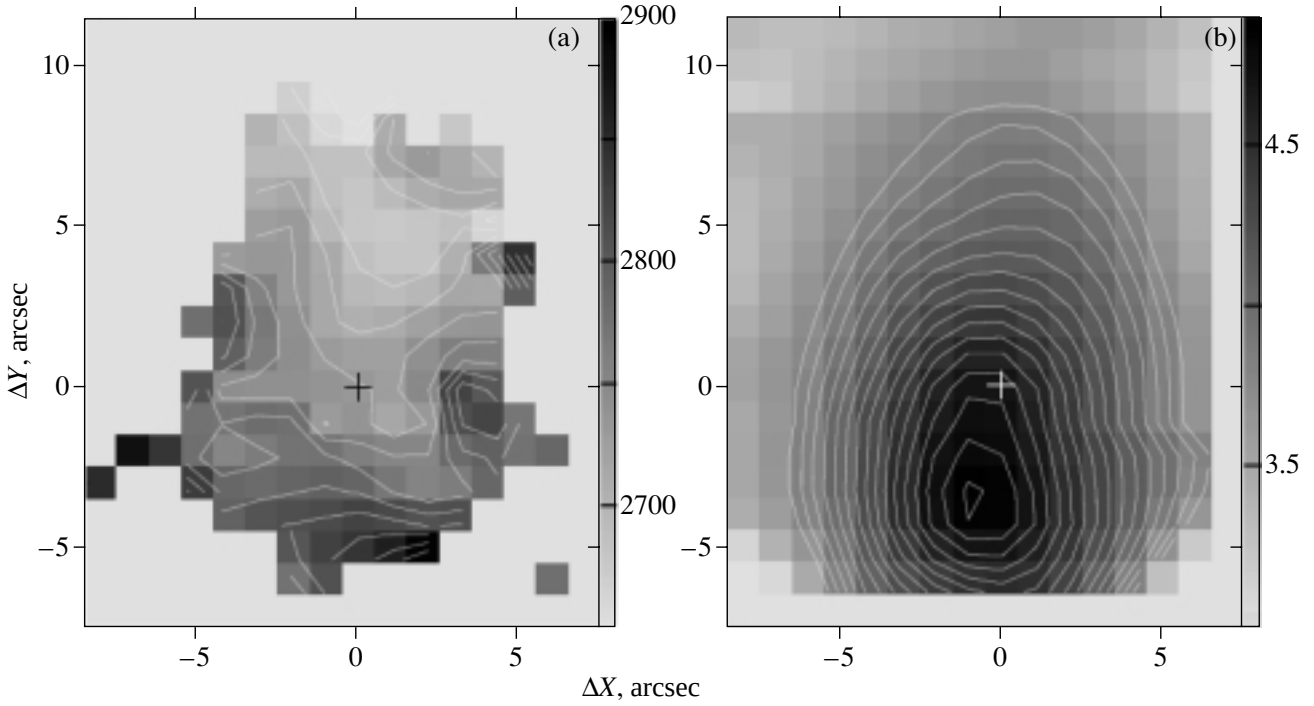

Fig. 3. (a) The stellar radial-velocity field and (b) the continuum-flux distribution in the range 5200–5450 Å; the cross marks the position of the kinematic center.

Combes (1994) concluded that counterrotation was observed in the central part of the galaxy ($|r| \leq 5''$). However, a detailed photometric study of this galaxy (Karataeva *et al.* 2001) shows that we see the total radiation from the galactic disk and inner ring exactly at distances of $4''$–$10''$ from the center. Therefore, the interpretation of the observed radial velocities in this range is rather complex and contradictory.

Where the slit crosses the spiral arms ($|r| \geq 20''$), the measurement errors of the radial velocities are large because the lines are weak. However, the velocity relative to the system center is, on average, $\sim$30 km s^{-1}, with the southern side approaching us and the northern side receding.

The radial-velocity curve along the ring major axis (Fig. 2b) is also complex in shape. This is most likely because the cut passes through different galactic structures.

Thus, we see that despite a wealth of information obtained with the long-slit spectrograph, the interpretation of the observed radial velocities of the emitting gas is ambiguous. One-dimensional cuts are not enough to understand the kinematics of such a multicomponent object. It is necessary to investigate the two-dimensional velocity fields of the gas and stars. This is the goal of our MPFS observations.

KINEMATICS OF THE STELLAR AND GASEOUS COMPONENTS

Data on the MPFS observations are also given in the table. Below, we discuss the results of our 2D spectroscopy.

The Radial-Velocity Distribution for the Stellar Component

We determined the radial velocities of the stellar component by cross-correlation analysis. Since the stellar-velocity dispersion turned out to be smaller than the spectrograph instrumental profile ($\sigma \leq$ 70 km s^{1-}), we could not reliably measure it and study its variations across the galaxy. The low stellar-velocity dispersion confirms the conclusion of Karataeva *et al.* (2001) that UGC 5600 is a late-type galaxy.

Figure 3 shows the stellar radial-velocity field and the continuum (5200–5450 Å) intensity distribution. We see from Fig. 3a that the isovels are complex in shape. The isophotes in the continuum image (Fig. 3b) are clearly distorted. These distortions may be due to the presence of a feature in the nuclear region; to the clumpy structures in the ring superimposed on the galaxy main body; due to the nonuniform distribution of dust, whose presence follows from the IR fluxes (Richter *et al.* 1994), both in the ring and in the galactic disk. The presence of dust may also affect the pattern of the radial-velocity field.

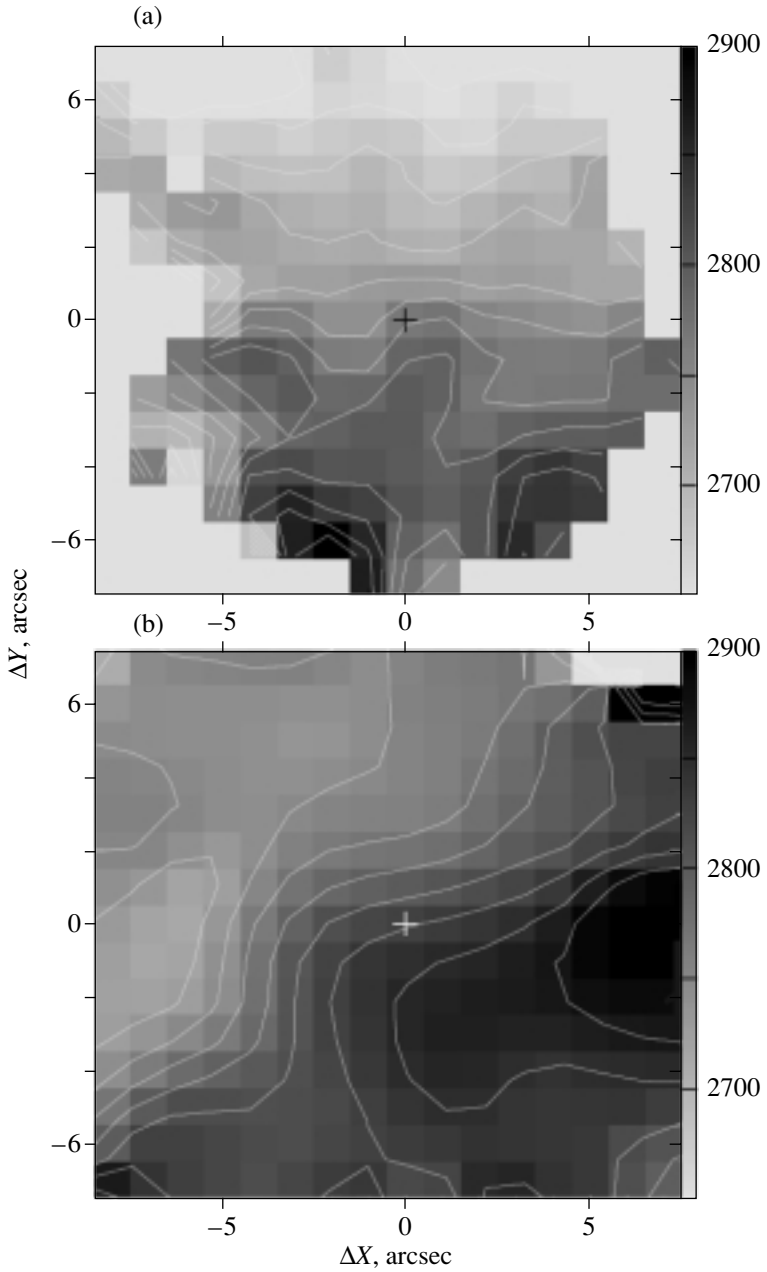

Fig. 4. The radial-velocity fields for the galaxy central region: (a) for stars and (b) for gas, as constructed from Hα at 1″ seeing.

If we consider the central region belonging to the galactic disk, then the shape of the isovels mainly corresponds to the circular rotation of stars around the galaxy minor axis. For this region, we constructed the average rotation curve and the radial dependence of the kinematic-axis position angle. We used the tilted-ring technique (Begeman 1989; Moiseev and Mustsevoi 2000): the velocity field is broken down into elliptical rings of fixed width and the rotation velocity $V(r)$ and the kinematic-axis position angle $PA(r)$ are determined in each ring under the assumption of circular rotation. In addition, conclusions about the pattern of noncircular motions can be drawn from an analysis of the variations in the position of the kinematic axis and in the disk inclination to the line of sight.

The cross in Fig. 3 marks the position of the kinematic center of the stellar component, which was determined from the symmetry of the velocity field. The photometric and kinematic centers do not coincide, but the separation between them is less than 2″. The differences in the positions of the photometric and kinematic centers may be due to the presence of features in the circumnuclear region. Within the

accuracy of our modeling, the position angle of the kinematic axis $PA_{kin} = 2°$ coincides with the position angle of the photometric axis from Karataeva *et al.* (2001) and the system heliocentric velocity is 2740 ± 15 km s^{-1}. The system velocities that we determined from stars and gas differ approximately by 50 km s^{-1}; such differences are also observed for other galaxies (see, e.g., Sil'chenko 1998). The inclination of the stellar disk to the plane of the sky is $\sim 50°$, which is close to the value from Karataeva *et al.* (2001), and its rotation velocity slowly increases from 40 km s^{-1} (at a distance of $2''$) to 90 km s^{-1} (at $8''$).

When analyzing in detail the stellar radial-velocity field in the circumnuclear region, which is $\sim 2''$ in size, we suspected a peculiarity in the behavior of the isovels. However, the space resolution was too low to study it. Therefore, as we pointed out above, additional observations were carried out at seeing no worse than $1''$. Figure 4a shows the stellar radial-velocity field constructed from these data. Noncircular stellar motions clearly show up in a $\sim 2''$ region, which may suggest the existence of a minibar that is possibly formed in the disk because of the close passage of a companion (Noguchi 1987). The nearly triangular shape of the isophotes in Fig. 3b and the asymmetry in the circumnuclear region of the photometric cuts along the galaxy major axis shown in Fig. 2 from Karataeva *et al.* (2001) are also consistent with this suggestion. However, observations with a higher space resolution are needed for the final conclusion about the presence of a minibar to be reached.

Two features that recede with velocities ~ 40–60 km s^{-1} show up at $\sim 4''.5$ from the center in the NE and W directions (Fig. 3a). The W feature may belong to the inner ring. At the same time, the velocity of the NE feature is opposite to the velocity of the ionized gas in the inner ring, while this feature is located far (~ 2 kpc) from the disk center.

The Radial-Velocity Distribution for the Gaseous Component

We constructed the radial-velocity fields of the gas from hydrogen (Hα, Hβ) lines and from the forbidden [O III] λ5007 line. They all proved to be similar. The radial-velocity curves obtained from the velocity fields are in good agreement with the radial-velocity curves given in the preceding section.

Hα is the brightest emission line. Since the accuracy of measuring the radial velocities from it is higher, we present below only the data obtained from this line. For convenience of comparing the behaviors of the gaseous and stellar components, Fig. 4 shows the velocity fields of the stars (a) and the ionized gas (b) for the galaxy central region. A detailed analysis of the gas kinematics indicates that the slope of the isovels in the central part, $\sim 3''$ in diameter, is identical to the slope of the isovels for the stellar component. Further out, in the E–W direction, the isovels are turned through about 90°, suggesting gas rotation about the galaxy major axis.

To analyze the behavior of the emitting gas in the galactic inner ring, Fig. 5 shows the Hα intensity distribution, the total radial-velocity field, and continuum (6100–6300 Å) intensity variations. In contrast to the continuum intensity distribution (Fig. 5c), the Hα image (Fig. 5a) is elongated from east to west and coincides with the ring location, but the ellipticity of the Hα isophotes changes. They become rounder as one approaches the center. South of the center, the isophotes flatten. This isophotal behavior can be explained by assuming that there are two gaseous components. The first is associated with the galactic disk, and the second is associated with the inner ring. In the central region, we observe the total radiation from the two components. The isophote flattening south of the center may imply that on this side, the ring is projected onto the galactic disk, and its dust partially absorbs the disk radiation. Note that our assumption is in conflict with the conclusion of Karataeva *et al.* (2001). Since the region with blue color indices on the southern side is narrower than on the northern side, these authors assumed that the ring was projected onto the galactic body north of the center and was seen through it south of the center. However, such a peculiarity of the color indices may be due to the nonuniform ring structure. For example, it may stem from the fact that the southern side of the ring is slightly narrower than its northern side. Therefore, the behavior of the Hα isophotes seems to characterize the ring orientation more reliably, especially since the isophotes are similar in shape in all emission lines.

The existence of two gaseous components must affect the shape of the line profiles along the galaxy major axis, particularly at distances of $5''$–$10''$ south and north of the center, where the ring and the disk are superimposed on each other. A significant scatter of points is observed precisely in these segments of the radial-velocity curve (Fig. 2a). The line profiles are irregular in shape and can be fitted by two Gaussians. However, the reliability of this fit is low, because the errors are large.

The assumption that there are two gaseous components is also confirmed by the shape of the isovels in Figs. 4b and 5b. At the center, where the directions of the isovels for the gaseous and stellar components coincide, the radial-velocity field is determined by the motion of the gas belonging to the galactic disk, while, starting approximately from $2''$ east and west of the center, the ring gas motion shows up clearly.

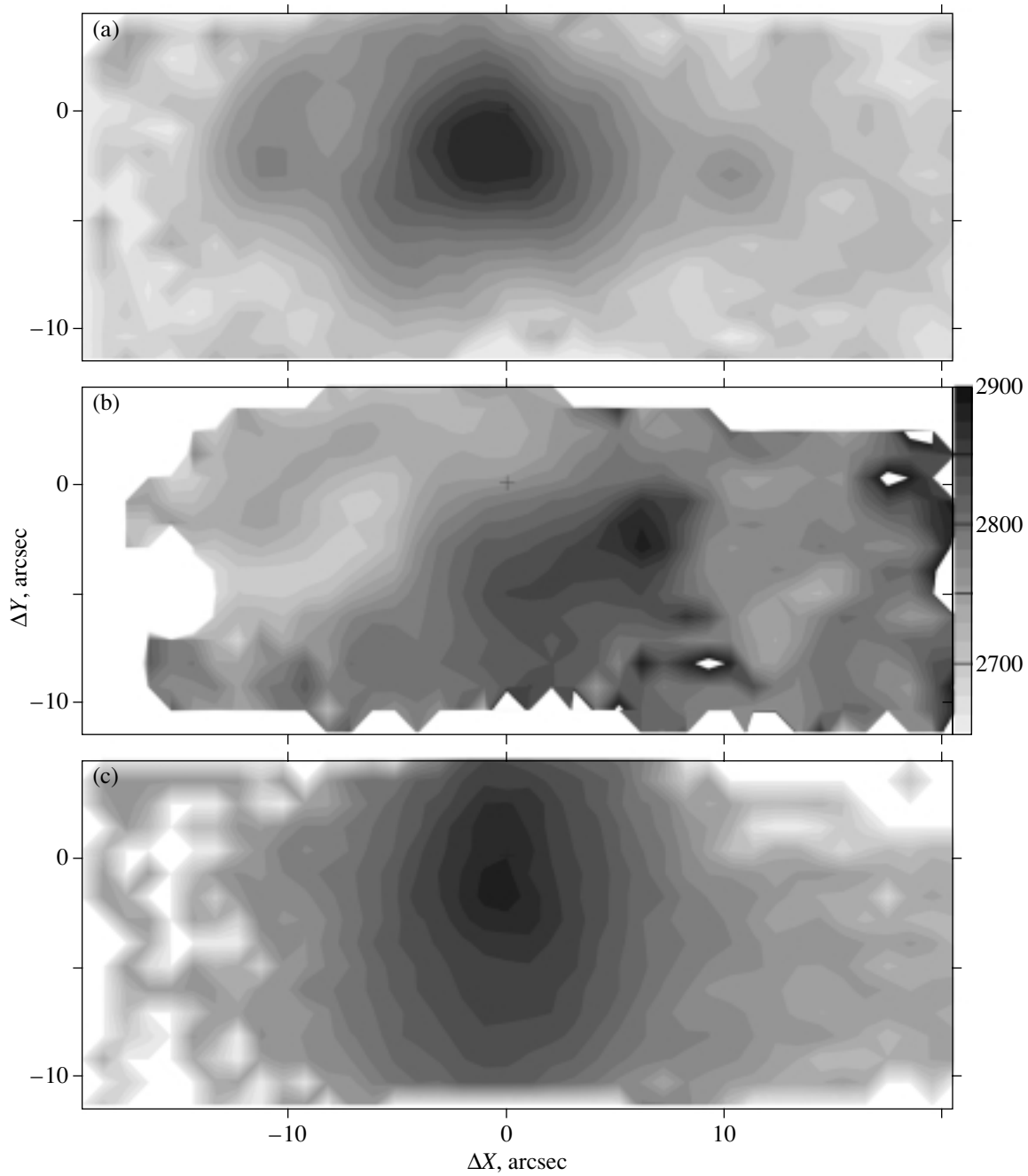

Fig. 5. (a) The intensity distribution, (b) the radial-velocity field in Hα, and (c) the continuum intensity distribution in the range 6100–6300 Å (the intensities are on a logarithmic scale).

Assuming circular rotation, we attempted to model the gas motion in the ring (without the central region $\sim 3''$ in size) and to estimate the ring inclination to the plane of the sky. We estimated the latter to be $\sim 70°-75°$. Knowing the inclination of the galactic disk and the ring to the plane of the sky, we can determine the angle between the disk and the ring from the relation

$$\cos \Delta i = \pm \sin i_1 \sin i_2 \cos (\mathrm{P}A_1 - \mathrm{PA}_2) \quad (1)$$
$$+ \cos i_1 \cos i_2,$$

where i_1 and i_2 are the disk and ring inclinations to the plane of the sky, $\mathrm{P}A_1$ and $\mathrm{P}A_2$ are the position angles of the major axes of the galactic disk and the ring. This angle was found to be about $78° \pm 5°$; i.e., the ring is polar.

A close examination of Fig. 5b reveals several features in the radial-velocity field. For example, there is a feature receding with a velocity of ~ 100 km s^{-1} at

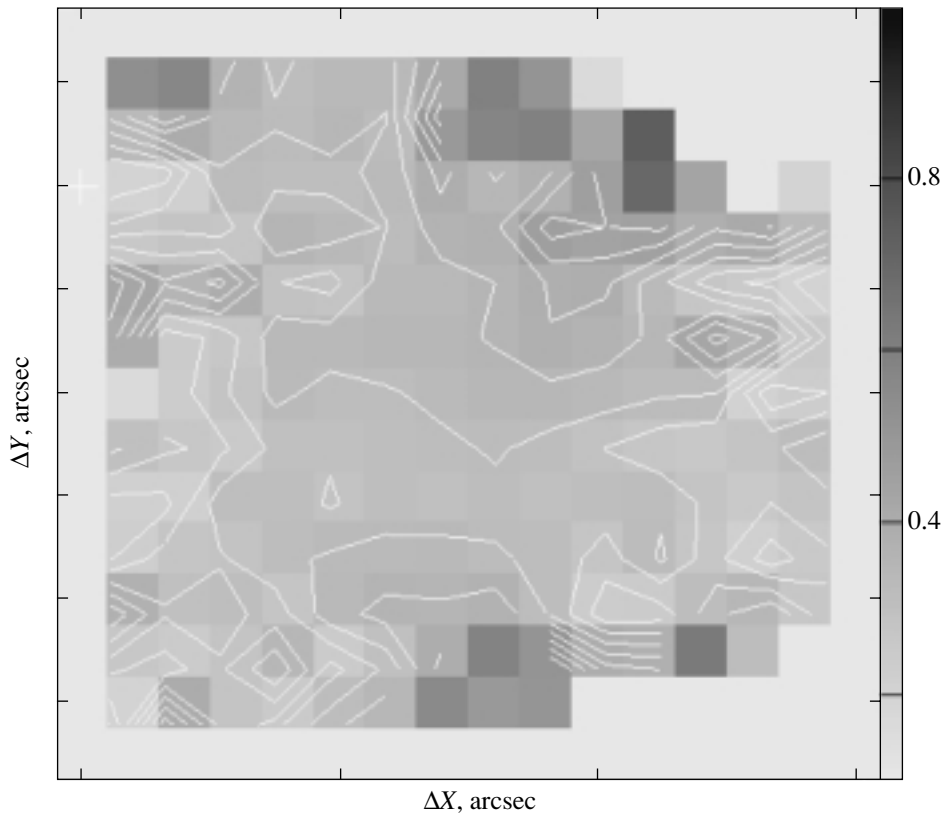

Fig. 6. The distribution of the [N II]/Hα intensity ratio in the galaxy $14'' \times 13''$ central region.

a distance of $7''$ to the west. Such features may result from the nonuniform ring structure. Their presence may stem from the fact that individual bright H II regions fall on the line of sight.

The right part of Fig. 5b, starting from $17''$–$20''$, corresponds to the spiral arm receding at a velocity of $\sim$30–40 km s^{-1}.

CHARACTERISTICS OF THE EMISSION-LINE REGIONS

We measured the Hα FWHMs and the relative emission-line intensities at various distances from the nucleus. Our measurements for the nuclear region closely agree, within the error limits, with the results of Reshetnikov and Combes (1994), who showed that the nuclear emission originates from H II regions. Figure 6 presents the distribution of the [N II]/Hα ratio for the central region of the galaxy. The logarithms of this ratio fall within the range -0.6 to -0.35, which is characteristic of H II regions (Veilleux and Osterbrock 1987). Therefore, we may conclude that the emission lines in the nucleus and in the ring originate in H II regions. The intensity ratio of forbidden and permitted lines is virtually constant along the major axis of the polar ring, suggesting that the physical conditions in the emission-line regions are similar. The forbidden-line intensity increases with distance from the nucleus along the galaxy major axis compared to Hα. To confirm that the increase in forbidden-line intensity (Fig. 6) is unrelated to the edge effects of the multipupil spectrograph, we considered the data acquired with the long-slit spectrograph. It turned out that the [N II]/Hα ratio at the center is 0.35 and then gradually increases, reaching 0.6 at $5''$–$7''$; further out, it is roughly constant up to $10''$–$12''$. The emission-line intensity decreases with distance from the center and starting from about $14''$, the signal-to-noise ratio is ≤ 3. Therefore, we cannot reliably determine the line intensity ratio in these regions.

The strengthening of the nitrogen forbidden lines compared to Hα appears to be due to an increase in the importance of the collisional excitation as the galactic gaseous disk interacts with the polar-ring gas.

Here, our prime objective was to study the kinematics of the gaseous and stellar components. Therefore, the observations were carried out only in the green and red spectral ranges, and we cannot determine the physical conditions in H II regions and their chemical composition from our data. However, we

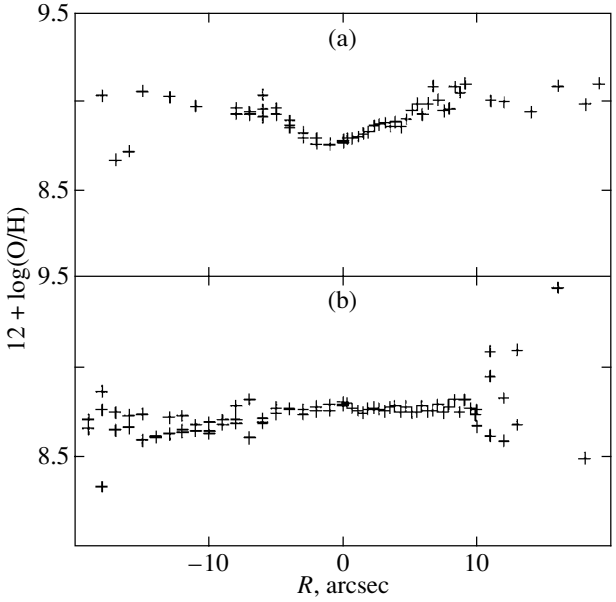

Fig. 7. Variations in the oxygen abundance (a) along the galaxy major axis and (b) along the major axis of the polar ring.

attempted to estimate the metallicity in the emission-line regions.

According to Denicolo *et al.* (2001), there is an empirical relation between log([N II] λ6583/Hα) and log(O/H):

$$12 + \log(\mathrm{O/H}) = 9.12(\pm 0.05) + 0.73(\pm 0.10) \times \log([\mathrm{N\ II}]\ \lambda 6583/\mathrm{H}\alpha); \quad (2)$$

its form depends neither on the flux calibration nor on the reddening corrections. Using this relation, we estimated the metallicity in H II regions. Figure 7 shows the radial distribution of $12 + \log(\mathrm{O/H})$ along the galaxy major axis and along the ring major axis. As was already pointed out above, the shape of the distribution in Fig. 7a results from an increase in the nitrogen-line intensity compared to the Hα intensity, possibly because the collisional excitation increases in importance. Further, the empirical relation (2) is reliable for an [N II]/Hα ratio below 0.5. Therefore, the increase in $12 + \log(\mathrm{O/H})$ in the regions between $5''$ and $10''$ north and south of the center may not result from an increase in metallicity. The values of $12 + \log(\mathrm{O/H})$ are virtually constant along the major axis of the polar ring (Fig. 7b), suggesting that the physical conditions in the ring are homogeneous.

The mean metallicity in the circumnuclear region and in the polar ring is $\approx$8.8, which corresponds to $0.9Z_\odot$. A similar estimate was also obtained from [S II] lines by using the sum of the [S II] λ6717 + λ6731 line fluxes. This result corresponds to normal evolution in galaxies with such luminosities (Richer *et al.* 1998).

The nearly solar metallicity in the polar ring implies that it cannot not be produced by the capture of a dwarf companion.

CONCLUSIONS

In conclusion, we summarize our main results.

(1) Based on 2D spectroscopy, we constructed the radial velocity fields of the stellar and gaseous components for the central regions of the peculiar galaxy UGC 5600.

(2) An analysis of these fields revealed two kinematic subsystems: the first is related to the galactic disk and the second is related to the inner ring.

(3) The angle between the disk and ring planes was found to be about $78^\circ \pm 5^\circ$; i.e., the inner ring is polar. This provides compelling evidence that the galaxy UGC 5600 belongs to PRGs.

(4) We established from the intensity ratio of forbidden and permitted lines that the emission originates in H II regions. The metallicity was estimated; it proved to be nearly solar, which rules out a dwarf galaxy as the donor in forming the polar ring.

ACKNOWLEDGMENTS

We are grateful to the 6-m telescope committee for allocating telescope observational time and to the SAO staff: V.L. Afanas'ev for assistance in the MFPS observations and for providing data reduction software; A.N. Burenkov for assistance in the UAGS observations and data reductions. We wish to thank O.K. Sil'chenko for helpful discussions and remarks. This study was supported in part by the "Astronomy" program (project no. 1.2.6.1) and the "Integration" program (project no. A0007).

REFERENCES

1. V. L. Afanasiev, A. N. Burenkov, V. V. Vlasyuk, and S. V. Drabek, Otchet SAO RAN, 234 (1995).
2. V. L. Afanasiev, S. N. Dodonov, and V. L. Moiseev, in *Stellar Dynamics: from Classic to Modern,* Ed. by L. P. Ossipkov and I. I. Nikiforov (St. Petersburg, 2001).
3. K. G. Begeman, Astron. Astrophys. **223**, 47 (1989).
4. G. Denicolo, R. Terlevich, and E. Terlevich, Mon. Not. R. Astron. Soc. **330**, 69 (2002).
5. G. Galletta, L. J. Sage, and L. S. Sparke, Mon. Not. R. Astron. Soc. **284**, 773 (1997).
6. G. M. Karataeva, V. A. Yakovleva, V. A. Hagen-Thorn, and O. V. Mikolaichuk, Pis'ma Astron. Zh. **27**, 94 (2001) [Astron. Lett. **27**, 74 (2001)].
7. A. V. Moiseev, Bull. SAO **51**, 11 (2001); astro-ph/0111219 (2001).

8. A. V. Moiseev and V. V. Mustsevoi, Pis'ma Astron. Zh. **26** (2000).
9. M. Noguchi, Mon. Not. R. Astron. Soc. **228**, 635 (1987).
10. V. P. Reshetnikov and F. Combes, Astron. Astrophys. **291**, 57 (1994).
11. M. G. Richer, M. L. McCall, and G. Stasinska, astro-ph/9809302 (1998).
12. O. R. Richter, P. D. Sackett, and L. S. Sparke, Astron. J. **107**, 99 (1994).
13. O. K. Sil'chenko, Astron. Astrophys. **330**, 412 (1998).
14. J. Tonry and M. Davis, Astron. J. **84**, 1511 (1979).
15. S. Veilleux and D. E. Osterbrock, Astrophys. J., Suppl. Ser. **63**, 295 (1987).
16. B. C. Whitmore, R. A. Lucas, D. B. McElroy, *et al.*, Astron. J. **100**, 1489 (1990).

Translated by N. Samus'

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 454–464. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 516–528.

Kinematic Parameters of Young Subsystems and the Galactic Rotation Curve

M. V. Zabolotskikh[1*], A. S. Rastorguev[1,2], and A. K. Dambis[2]

[1] *Moscow State University, Vorob'evy gory, Moscow, Russia*

[2] *Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119992 Russia*

Received February 1, 2002

Abstract—We analyze the space velocities of blue supergiants, long-period Cepheids, and young open star clusters (OSCs), as well as the H I and H II radial-velocity fields by the maximum-likelihood method. The distance scales of the objects are matched both by comparing the first derivatives of the angular velocity Ω' determined separately from radial velocities and proper motions and by the statistical-parallax method. The former method yields a short distance scale (for $R_0 = 7.5$ kpc, the assumed distances should be increased by 4%), whereas the latter method yields a long distance scale (for $R_0 = 8.5$ kpc, the assumed distances should be increased by 16%). We cannot choose between these two methods. Similarly, the distance scale of blue supergiants should be shortened by 9% and lengthened by 3%, respectively. The H II distance scale is matched with the distance scale of Cepheids and OSCs by comparing the derivatives Ω' determined for H II from radial velocities and for Cepheids and OSCs from space velocities. As a result, the distances to H II regions should be increased by 5% in the short distance scale. We constructed the Galactic rotation curve in the Galactocentric distance range 2–14 kpc from the radial velocities of all objects with allowance for the difference between the residual-velocity distributions. The axial ratio of the Cepheid+OSC velocity ellipsoid is well described by the Lindblad relation, while $\sigma_u \approx \sigma_\nu$ for gas. The following rotation-curve parameters were obtained: $\Omega_0 = (27.5 \pm 1.4)$ km s^{-1} kpc^{-1} and $A = (17.1 \pm 0.5)$ km s^{-1} kpc^{-1} for the short distance scale ($R_0 = 7.5$ kpc); and $\Omega_0 = (26.6 \pm 1.4)$ km s^{-1} kpc^{-1} and $A = (15.4 \pm 0.5)$ km s^{-1} kpc^{-1} for the long distance scale ($R_0 = 8.5$ kpc). We propose a new method for determining the angular velocity Ω_0 from stellar radial velocities alone by using the Lindblad relation. Good agreement between the inferred Ω_0 and our calculations based on space velocities suggests that the Lindblad relation holds throughout the entire sample volume. Our analysis of the heliocentric velocities for samples of young objects reveals noticeable streaming motions (with a velocity lag of $\sim$7 km s^{-1} relative to the LSR), whereas a direct computation of the perturbation amplitudes in terms of the linear density-wave theory yields a small amplitude for the tangential perturbations. © *2002 MAIK "Nauka/Interperiodica".*

Key words: *Galactic kinematics, rotation curve, distance scale*

INTRODUCTION

The study of the kinematics of Galactic subsystems remains one of the most important fields of Galactic astronomy. The parameters of the Galactic rotation curve were determined repeatedly from HI and HII data (Clemens 1985; Fich *et al.* 1989; Merrifield 1992; Brandt and Blitz 1993; Nikiforov and Petrovskaya 1994; Honma and Sofue 1997; Nikiforov 1999) and stellar radial velocities (Karimova and Pavlovskaya 1973; Pont *et al.* 1994; Dambis *et al.* 1995; Glushkova *et al.* 1998). High-precision proper motions and trigonometric parallaxes that became available with the release of the HIPPARCOS catalog (The HIPARCOS and TYCHO catalogs, ESA SP-1200, 1997) stimulated further works aimed at refining the angular velocity Ω_0 and the form of the rotation curve in the local solar neighborhood (Feast *et al.* 1998; Rastorguev *et al.* 1999; Dambis *et al.* 2001). It should be pointed out that the reliability of the resulting rotation curves depends first and foremost on the correctness of the adopted distance scale of objects under study. Objects with known distances—classical Cepheids, open star clusters (OSC), and OB-associations—allow the rotation curve to be determined only out to heliocentric distances of 4–5 kpc, whereas H I and H II kinematic data allow constructing the rotation curve over a considerably wider interval of Galactocentric distances. The main problem is that the distances of giant molecular clouds (GMC) and, consequently, those of H II-regions, are determined from their single hot exciting stars whose distance scale is prone not

*E-mail: zabolot@lnfm1.sai.msu.ru

only to random but also to systematic errors. In this paper we matched for the first time the distance scale of GMC to the most accurately determined (in the random and systematical sense) distances, i.e., those of long-period classical Cepheids and OSC, and inferred the kinematic parameters using an algorithm that allows for initial-data errors, for the ellipsoidal distribution of residual velocities, and for the errors of systemic radial velocities that result from the propagation of distance errors (Rastorguev 2001).

Our second task was to compute the parameters of the spiral pattern of the Galaxy. Selection effects, i.e., the incompleteness of the sample due to interstellar extinction, makes it very difficult to localize spiral waves by analyzing only the space distribution of young objects. The use of kinematic data appears to be a more promising approach, because it is insensitive to selection effects (Mishurov *et al.* 1979). Thus Mishurov *et al.* (1997) determined, by analyzing the radial velocities of classical Cepheids exclusively, the principal parameters of the spiral pattern including the velocity-field perturbation amplitudes and concluded that the Sun is located near the corotation circle. The perturbations due to the spiral density wave are comparable in magnitude to the velocity dispersion of young subsystems. Therefore, only after the release of the HIPPARCOS and TYCHO-2 catalogs making high-precision proper motions available did it become possible to analyze the space velocity field of young objects. Torra *et al.* (2000) used the radial velocities of OB-stars and Cepheids and HIPPARCOS proper motions of these objects to infer a pattern speed of $\Omega_P \approx (31 \pm 4)$ km s^{-1} kpc^{-1}. Lepine *et al.* (2001) also concluded that the Sun is near the corotation circle by assuming a superposition of a two- and four-armed pattern. At the same time, Rastorguev *et al.* (2001) concluded that the Sun is inside the corotation circle by analyzing long-period Cepheids and young OSC, and Mel'nik *et al.* (2001) came to the same conclusion based on their study of the pattern of systematical noncircular motions of OB-associations. Here we explore this issue further.

OBSERVATIONAL DATA

We used young OSC and long-period Cepheids as a reference sample for matching the distance scales of various objects. Our reference sample included 89 young OSC with $\log T < 7.6$ and heliocentric distances determined by Dambis (1999) by fitting Kholopov's (1980) ZAMS with an allowance for evolutionary deviations based on Geneva-group isochrones (Maeder and Meynet 1991). The radial velocities of cluster members were determined by Glushkova based on published data and can be found in the paper by Rastorguev *et al.* (1999). The proper motions of clusters were computed from those of their member stars found in the HIPPARCOS catalog (Baumgardt *et al.* 2000).

Our reference sample included 113 classical Cepheids with periods $P > 9^{\rm d}$ (or ages $\log T < 7.6$ as implied by the period-age relation of Efremov (1989)) and heliocentric distances computed using the fundamental-mode period-luminosity relation of Berdnikov *et al.* (1996):

$$\langle M_K \rangle_I = -5.46^m - 3.52^m \log P$$

in accordance with the procedure described therein. An earlier statistical-parallax analysis (Rastorguev *et al.* 1999) showed that the sample of Cepheids with shorter periods is not homogeneous in terms of pulsation mode and may be contaminated by first-overtone pulsators. We used published Cepheid radial velocities and HIPPARCOS proper motions. Young OSC and long-period Cepheids make up a kinematically homogeneous sample consisting of 176 and 142 objects with radial velocities and proper motions, respectively, including 124 objects with space velocities.

We performed a separate analysis of a blue-supergiant sample consisting of 102 stars with heliocentric distances tied to the OSC distance scale (Dambis 1990). The kinematic data for these stars were compiled by A.K. Dambis with the proper motions adopted from the HIPPARCOS catalog, and radial velocities, from the catalogs of Barbier-Brossat and Figon (2000) and Wilson–Evans–Batten (WEB) (Duflot *et al.* 1995).

Brandt and Blitz (1993) published the distances and radial velocities for a total of 206 H II-regions. We selected 203 of these objects with spectroscopic or photometric distances inferred from their exciting stars. The radial velocities of H II-regions were determined from the CO (2.2.-mm) radio lines of their associated molecular clouds. We did not include three H II-regions in the final list because of their large residual velocities relative to the provisional rotation-curve solution. The catalog mentioned above also gives standard errors of individual distance and radial velocities.

We adopted 150 tangent-point radial velocities of H I clouds from Fich *et al.* (1989). Note that published H I and H II radial velocities are traditionally corrected for the solar motion relative to the standard apex assumed to coincide with the local standard of rest (LSR), and we therefore first converted them into heliocentric radial velocities.

METHOD OF ANALYSIS

We used the techniques of maximum-likelihood and statistical parallax (including its simplified version) to compute the kinematic parameters and refine the distance scales involved. See Murray (1986) for a description of the principal ideas of the statistical-parallax method used in this paper. The tangential velocity of a star is computed from its proper motion and distance and therefore depends on the adopted distance scale, whereas radial velocities are distance independent. The essence of the method is to reconcile the fields of radial and tangential velocities in terms of some model of the field of systematic motions and ellipsoidal distribution of residual velocities. A number of authors applied this method with success. Hawley *et al.* (1986); Popovski and Gould (1998), Gould and Popowski (1998), Fernley *et al.* (1998), Popowski (1998), Fernley *et al.* (1998), Tsujimoto *et al.* (1998), and Dambis and Rastorguev (2001) used it to refine the distance scale of RR Lyrae type variables. In our previous paper (Rastorguev *et al.* 1999) we applied this method for the first time, albeit in a somewhat simplified form ignoring the scatter of absolute magnitudes, to analyze the space velocity field of young objects of the Galactic disk, which are characterized by small residual velocity dispersions.

In this paper we also apply a simplified version of the statistical-parallax technique (as used, e.g., by Feast *et al.* 1998), based on reconciling the kinematic parameters inferred separately from radial velocities and proper motions. Thus, it is well known that Oort's constant A inferred from proper motions is much less sensitive to the adopted distance scale than is the value of the same constant inferred from radial velocities. This allows not only the kinematic parameters to be determined but also the distance scale of objects under study to be refined.

Consider now a model of the field of space velocities that includes both differential rotation and effects due to a spiral density wave. The residual velocity of a star can be written in the form of the following column vector:

$$\Delta\mathbf{V} = \mathbf{V}_{\rm obs} - \mathbf{V}_{\rm sys} = \mathbf{V}_{\rm obs} - \mathbf{V}_{\rm sun} - \mathbf{V}_{\rm rot} - \mathbf{V}_{\rm spir},$$

where $\mathbf{V}_{\rm obs}$ is the observed space velocity; $\mathbf{V}_{\rm sys}$, the total velocity of systematic motions including: $\mathbf{V}_{\rm sun}$, the mean heliocentric velocity of the sample studied; $\mathbf{V}_{\rm rot}$, the contribution of Galactic differential rotation; and $\mathbf{V}_{\rm spir}$, the perturbation due to the spiral density wave. To allow for spiral-pattern effects, we used a very simple kinematic model based on linear density-wave theory by Lin *et al.* (1969) with the perturbation of potential in the form of a running wave:

$$\Phi_S = A_\Phi \cos\chi,$$

where $A_\Phi < 0$ is the amplitude of perturbations and

$$\chi = m(-\theta + \cot i \ln(R/R_0)) + \chi_0,$$

the phase angle of the object in the wave (it increases toward the Galactic center). Here m is the number of arms; θ, the position angle of the object (measured in the direction of rotation); i, the pitch angle of spiral arms ($i < 0$ for trailing spirals); χ_0, the phase angle of the Sun; and R and R_0, the Galactocentric distances of the Sun and the object, respectively. The radial V_R (which in the arm is directed toward the Galactic center) and azimuthal V_θ (directed along differential rotation at the outer edge of the arm) components of velocity perturbation can be written in the following form:

$$V_R = f_R \cos\chi, \quad V_\theta = f_\theta \sin\chi,$$

where f_R and f_θ are the amplitudes of velocity perturbations (Rohlfs 1977):

$$f_R = \frac{kA_\Phi}{\kappa}\frac{\nu}{1-\nu^2}F_\nu^{(1)}(x),$$

$$f_\theta = -\frac{kA_\Phi}{2\Omega}\frac{1}{1-\nu^2}F_\nu^{(2)}(x).$$

We now use standard designations:

$$k = \frac{m\cot i}{R}, \quad \kappa = 2\Omega\sqrt{1-\frac{A}{\Omega}},$$

$$x = \left(\frac{k\sigma_u}{\kappa}\right)^2, \quad \nu = \frac{m(\Omega_P - \Omega)}{\kappa}.$$

Here, k is the radial wavenumber; κ, the epicyclic frequency; A, Oort's constant; $F_\nu^{(1)}(x)$ and $F_\nu^{(2)}(x)$, the reduction factors; x, the Toomre instability parameter; σ_u, the dispersion of radial velocities; ν, the relative frequency with which the object rotating in a circular orbit meets a passing spiral wave; Ω, the angular velocity of differential Galactic rotation; and Ω_P, the angular velocity of the rigid rotation of the spiral pattern (i.e., the pattern speed).

Residual space velocities are usually assumed to have a three- dimensional normal distribution:

$$f(\Delta\mathbf{V}) = (2\pi)^{-3/2}|L_{\rm obs}|^{-1/2} \times \exp\left\{-0.5\Delta\mathbf{V}^T \times L_{\rm obs}^{-1} \times \Delta\mathbf{V}\right\},$$

where $L_{\rm obs}$ is the matrix of covariances. The covariance matrix in our previous paper (Rastorguev *et al.* 1999) included only the ellipsoidal velocity distribution and the errors of radial velocities and proper motions, which is quite a justiable approach in the case of small errors in the adopted distances. The latter are related to the dispersion of absolute magnitudes as follows:

$$\sigma_M^2 = 4.71\langle(\delta r/r)^2\rangle.$$

Here, angular braces mean averaging over the distribution of distance errors. The covariance matrix can be easily shown (Rastorguev 2001) to have the following form:

$$L_{\rm obs} = L_{{\rm loc},e} + L_{\rm err} + 0.21\sigma_M^2 r_t^2 (P \times L_1 \times P') - 0.21\sigma_M^2 p r_t (M \times L_2 \times P') + 0.21\sigma_M^2 p^2 (M \times L_3 \times M'),$$

where matrices L_1, L_2, L_3, M, and P are equal to

$$L_1 = \frac{d\mathbf{V}_{\rm sys}}{dr_t} \times \frac{d\mathbf{V}_{\rm sys}^{\rm T}}{dr_t}, \quad L_2 = 2\mathbf{V}_{\rm sys} \times \frac{d\mathbf{V}_{\rm sys}^{\rm T}}{dr_t},$$
$$L_3 = G_S \times L_{S,0} \times G_S^{\rm T} + \mathbf{V}_{\rm sys} \times \mathbf{V}_{\rm sys}^{\rm T},$$

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix},$$

respectively, and the formulas for matrices G_S, $L_{{\rm loc},e}$, $L_{\rm err}$, and $L_{S,0}$ can be found in the paper by Rastorguev *et al.* (1999). Here p is the distance-scale factor defined as:

$$p = r_0/r_t,$$

where r_0 and r_t are the adopted (usually photometric) and true distance, respectively.

We inferred the unknown parameters including the scale factor p using the maximum-likelihood method, i.e., by minimizing the following function with summation taken over all objects of the sample under study:

$$\mathrm{LF} = -\sum_{i=1}^{N} \ln f(\Delta\mathbf{V}).$$

When refining the distance scale by reconciling the values of Oort's constant A, we set $p = 1$. We computed the parameter errors using the method proposed by Hawley *et al.* (1986).

RESULTS AND DISCUSSION

Kinematics of the Sample of Long-Period Cepheids and OSC

Our main task was twofold: to refine the distance scale of objects considered and to construct the rotation curve of the corresponding subsystem. We first applied the maximum-likelihood method to our sample of Cepheids and OSC with heliocentric distances $r < 4$ kpc and ignored spiral-pattern effects in the velocity field. Because the eventual correlation between the solar Galactocentric distance R_0 and the distance- scale factor prevents simultaneous determination of these parameters, and because of the uncertainty in the determination of R_0, we performed our computations twice with the two most commonly adopted values: $R_0 = 7.5$ and 8.5 kpc. We determined the angular velocity of Galactic rotation Ω_0 from space velocities of Cepheids and OSC and then used it to construct the Galactic rotation curve based on the radial velocities of all objects considered.

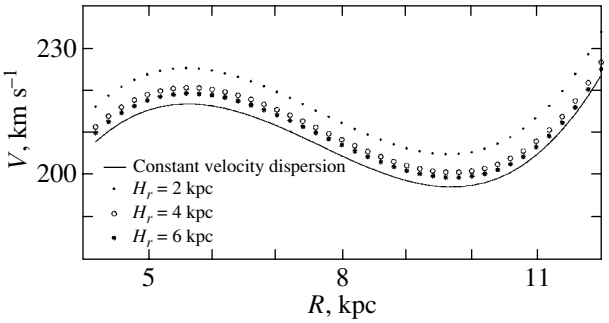

Fig. 1. Galactic rotation curves inferred with different scale lengths H_r and assuming constant velocity dispersion.

How distance errors affect the results. To elucidate the effect of the distance errors on the results obtained, we repeated our computations with three different standard errors of absolute-magnitude calibration: $0\overset{m}{.}1$, $0\overset{m}{.}15$, and $0\overset{m}{.}2$ for Cepheids and OSC. We set $R_0 = 7.5$ kpc and $p = 1$ in all three cases. The results are listed in Table 1. The columns of this table give the standard error of the absolute magnitude; heliocentric velocity components of the sample; velocity-ellipsoid axes; and rotation-curve parameters. The inferred kinematic parameters can be seen to be virtually independent of the adopted σ_M, and therefore in the following computations we used a compromise value of $\sigma_M = 0\overset{m}{.}15$, which agrees with the scatter of the period-luminosity relation for the Cepheids members of open clusters (Berdnikov *et al.* 1996).

The effect of the variation of velocity dispersion with galactocentric distance. The study of the kinematics and space distribution of objects in the disks of other galaxies showed that the disk surface brightness and velocity dispersion decrease exponentially with galactocentric distance, and the squared velocity dispersion is proportional to the surface density (van den Kruit and Freeman 1986; Bottema 1993). The corresponding scale length for our Galaxy can be estimated only indirectly and is most likely confined between 2 and 6 kpc depending on the age of the subsystem studied (Lewis and Freeman 1989; Kent *et al.* 1991; Malhotra 1995; Dehnen and Binney 1998; Freudenreich 1998; Drimmel and Spergel 2001). Let us assume that radial velocity

Table 1. Kinematic parameters of the Cepheid + OSC sample inferred adopting different standard errors of absolute magnitudes

σ_M	u_0	ν_0	w_0	σ_u	σ_ν	σ_w	Ω_0, km s^{-1} kpc^{-1}	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
	km s^{-1}								
$0^m.10$	−6.34	−12.39	−6.95	12.90	8.02	7.21	27.71	−4.66	1.18
$0^m.15$	−6.31	−12.33	−6.96	12.82	7.93	7.19	27.71	−4.66	1.17
$0^m.20$	−6.27	−12.25	−6.98	12.72	7.81	7.15	27.71	−4.65	1.15

Table 2. Kinematic parameters of the Cepheid + OSC sample inferred with different scale lengths of the assumed exponential radial decrease of velocity dispersions

H_r, kpc	u_0,	ν_0,	w_0,	σ_u,	σ_ν,	σ_w,	Ω_0, km s^{-1} kpc^{-1}	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
	km s^{-1}								
2	−6.77	−12.45	−6.94	14.40	7.88	7.15	28.76	−4.84	1.23
4	−6.35	−12.34	−6.95	13.25	7.93	7.17	28.22	−4.74	1.18
6	−6.30	−12.32	−6.95	13.04	7.94	7.18	28.05	−4.71	1.17

Table 3. Kinematic parameters and the distance-scale factor for the Cepheid + OSC sample inferred via statistical parallaxes

R_0, kpc	p	u_0,	ν_0,	w_0,	σ_u,	σ_ν,	σ_w,	Ω_0, km s^{-1} kpc^{-1}	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
		km s^{-1}								
7.5	0.86	−7.24	−11.51	−8.06	13.70	8.03	8.55	26.93	−4.27	0.94
8.5	0.84	−7.21	−12.33	−8.24	13.65	8.15	8.76	26.61	−3.66	0.73
Standard errors	±0.05	±2.10	±1.76	±1.61	±1.62	±1.18	±1.68	±1.35	±0.24	±0.19

dispersion varies exponentially with Galactocentric radius:

$$\sigma_u = \sigma_u^0 \exp\left(-\frac{R-R_0}{2H_r}\right),$$

Table 4. The first derivative of angular velocity inferred separately from radial velocities V_r and proper motions μ of the Cepheid + OSC sample and the resulting distance-scale factor

Method	R_0, kpc	$\Omega'(V_r)$, km s^{-1} kpc^{-2}	$\Omega'(\mu)$, km s^{-1} kpc^{-2}	p
1	7.5	−4.67	−4.53	0.97
2	7.5	−4.68	−4.42	0.94
1	8.5	−4.04	−3.97	0.98
2	8.5	−4.06	−3.88	0.96
Standard errors		±0.26	±0.34	±0.09

where σ_u^0 is the radial velocity dispersion in the solar neighborhood and H_r, the disk scale length parameter. As is evident from our Table 1 (see also results of Rastorguev *et al.* (1999) and Dehnen and Binney (1998)), in the neighborhood of the Sun the components of the velocity dispersion tensors of both the classical Cepheids + OSC and local MS-star sample obey the following Lindblad relation to a good accuracy:

$$\sigma_\nu = \sigma_u \frac{\kappa}{2\Omega}.$$

It is possible, assuming that this relation is obeyed at every point of the disk for the current values of angular rotation velocity and epicyclic frequency, to determine how the inferred kinematic parameters depend on the adopted disk scale length. In this analysis we can neglect the effect of the variation of the vertical velocity dispersion σ_w with Galactocentric distance, because, first, it is insignificant compared to the errors of tangential velocities V_b, and, second,

the allowance for the dependence of vertical velocity dispersion on Galactocentric distance has virtually no effect on the results obtained. Table 2 presents the kinematic parameters computed with $R_0 = 7.5$ kpc, and $p = 1$. The tabulated velocity-ellipsoid axes refer to the solar neighborhood. Note that Ω_0 is sensitive to the adopted scale length parameter.

Figure 1 shows how the inferred rotation curve changes with the adopted scale length. Although the exact scale for our sample is unknown, an analysis of the results of Kent *et al.* (1991) leads us to conclude that young subsystems are characterized by a relatively shallower decrease of radial dispersion with Galactocentric distance. According to our results, the rotation curve for $H_r = 2$ kpc is 9 km s^{-1} higher than if computed for constant velocity dispersion. Dehnen and Binney (1998) inferred a scale length of $\sim$2–2.5 kpc for old main-sequence stars; Drimmel and Spergel (2001) found a scale length of $0.28R_0$ by analyzing COBE/DIRBE data (note that both old and young stars contribute to infrared radiation). Since no accurate data are available about the relation between surface brightness and velocity dispersion, hereafter we assume that velocity dispersion remains constant along Galactocentric radius.

Refining the distance scale. Table 3 lists the kinematic parameters inferred treating the distance-scale factor as an unknown parameter. The initial distances to Cepheids and OSC are on the short distance scale. As is evident from the table, the distance-scale factor depends only slightly on the adopted R_0. Judging by these results, the adopted distance scale should be increased by 14–16%.

Besides the rigorous method of statistical parallaxes, we also used its simplified version, which involves comparing the values of the first derivative of angular velocity Ω' inferred separately from radial velocities with proper motions. We determined the kinematic and rotation-curve parameters of the sample under study from independent maximum-likelihood solutions based on radial velocities and proper motions. It can be easily seen that radial velocities of stars of flat subsystems allow neither w_0 nor σ_w to be accurately constrained. We therefore inferred Ω' in two ways: (1) by computing the heliocentric space velocity components u_0, ν_0, w_0 and vertical velocity dispersion σ_w of the sample under study from space velocities, and then fixing these values in separate radial velocity and proper-motion solutions; and (2) by substituting u_0 and ν_0 inferred from radial velocities into the proper-motion solution and substituting w_0 and σ_w inferred from proper motions into the radial velocity solution. Table 4 lists the resulting Ω' and distance-scale factors $p = \Omega'(\mu)/\Omega'(V_r)$.

The resulting mean distance-scale factor for the Cepheid+OSC subsystem is equal to $p = 0.96$ for $R_0 = 7.5$ kpc (with $\Omega' = -4.50$ km s^{-1} kpc^{-2}) and $p = 0.97$ for $R_0 = 8.5$ kpc (with $\Omega' = -3.95$ km s^{-1} kpc^{-2}). Again, we note a weak dependence of the distance-scale factor on the adopted solar Galactocentric distance (see Table 5).

Noteworthy are (see Tables 3 and 5) systematic differences between the distance-scale factors given by the statistical-parallax technique ($\sim$0.86) and by its simplified modification ($\sim$0.96). We analyzed the problem for possible biases using numerical simulations. To this end, we used the real coordinates and initial distances to the objects of our sample and simulated their "true" space velocities based on the earlier determined values of kinematic and rotation-curve parameters. We then added normally distributed errors to the "true" distances and space velocities and redetermined the kinematic parameters and distance-scale factor using both the rigorous statistical-parallax technique (space velocities) and its simplified modification. We set velocity errors based on the typical errors of observational data and ellipsoidal distribution of residual velocities. One hundred numerical simulations yielded a mean distance-scale factor of $p = 1.00 \pm 0.05$ and $p = 1.01 \pm 0.07$ by making inferences from space velocities or by comparing the first derivatives of angular velocity, respectively. The possible distance-scale factors were confined to the (0.85–1.15) interval, with, on the average, correlated deviations of the two values from unity. We cannot unambiguously choose between the two approaches to the distance-scale refinement. Since it is logical to associate the short and long distance scales with $R_0 = 7.5$ kpc and $R_0 = 8.5$ kpc, respectively, hereafter we inferred the kinematic parameters assuming that the Cepheid+OSC distance-scale factors of $p = 0.96$ and 0.84 correspond to $R_0 = 7.5$ and 8.5 kpc, respectively. In support of this conclusion, we determined R_0 from space velocities with fixed p. Our analysis yielded $R_0 = (7.4 \pm 1.0)$ and (8.3 ± 1.0) kpc for $p = 0.96$ and 0.84, respectively. The large errors of the resulting R_0 are due to the small size of the data sample used.

Determination of Ω_0 from radial velocities using Lundblad's relation. Note that the fact that the velocity dispersions of Cepheids and OSC obey the Lindblad relation allows the angular velocity of rotation Ω_0 at the solar Galactocentric distance to be estimated independently from radial velocities exclusively. When computing the kinematic parameters, the idea is to set as unknown only the radial velocity dispersion σ_u^0 at the solar Galactocentric distance and to determine the ratio of velocity ellipsoid axes from the Lindblad relation while setting the angular velocity and its derivative equal to their local values for each

Table 5. Kinematic parameters of the Cepheid + OSC sample inferred from space velocities using the mean distance-scale factors (determined by comparing the first derivatives of angular velocity)

R_0, kpc	u_0	ν_0	w_0	σ_u	σ_ν	σ_w	Ω_0, km s^{-1} kpc^{-1}	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
	km s^{-1}								
7.5	−6.55	−12.11	−7.25	13.04	7.92	7.55	27.47	−4.54	1.09
8.5	−6.38	−12.98	−7.18	12.80	8.04	7.44	27.37	−3.99	0.90
Standard errors	±1.77	± 1.71	±1.24	±1.49	±1.10	±1.22	±1.39	±0.24	±0.19

Table 6. Kinematic parameters of the blue-supergiant sample

R_0, kpc	u_0	ν_0	w_0	σ_u	σ_ν	σ_w	Ω_0, km s^{-1} kpc^{-1}	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
	km s^{-1}								
7.5	−6.04	−10.92	−7.12	11.49	8.96	5.13	29.60	−4.76	0.89
8.5	−6.18	−11.33	−7.86	11.63	9.41	5.71	29.14	−4.00	0.60
Standard errors	±1.93	±1.73	±1.08	±1.42	±1.14	±1.03	±1.62	±0.32	±0.53

Table 7. Kinematic and rotation-curve parameters inferred from H II data

R_0, kpc	u_0	ν_0	σ_u	σ_ν	Ω', km s^{-1} kpc^{-2}	Ω'', km s^{-1} kpc^{-3}
	km s^{-1}					
7.5	−8.11	−14.88	6.70	6.91	−4.77	1.26
8.5	−7.92	−15.73	6.56	7.03	−4.08	1.01
Standard errors	±1.71	±1.19	±2.07	±1.14	±0.27	±0.22

object. The resulting likelihood function therefore depends explicitly on the unknown angular velocity Ω_0. We applied this method with fixed $w_0 = -7$ km s^{-1} and $\sigma_w = 7$ km s^{-1} to the radial velocities of Cepheids and OSC with $p = 0.96$ and $R_0 = 7.5$ kpc to obtain $\Omega_0 = (26.5 \pm 8.7)$ km s^{-1} kpc^{-1}. The large error of the inferred angular velocity is fully explained by the errors of the inferred velocity dispersions, which are equal to 1.2–1.7 km s^{-1} (see Table 5). The surprisingly good agreement between the angular velocity values inferred from space and radial velocities indicates that the Lindblad relation is obeyed accurately enough throughout the entire space region studied.

Kinematics of the Blue-Supergiant Sample

We applied the maximum-likelihood technique to a sample of 102 blue supergiants. The dispersion of the inferred absolute magnitudes for these stars is higher than for Cepheids and OSC and is equal to $\sigma_M \approx 0\overset{m}{.}38$ (Dambis 1990). We estimated the distance-scale factor using the two methods described above. The maximum-likelihood method applied to space velocities of stars yielded $p = 0.97 \pm 0.08$, whereas a comparison of the first derivatives of angular velocity determined separately from radial velocities and proper motions yielded $p = 1.09 \pm 0.16$ (error estimated approximately). Both results agree fairly well with the correction factor to the blue-supergiant distance scale ($p = 1.03 \pm 0.04$) inferred from practically the same sample by comparing photometric and HIPPARCOS trigonometric parallaxes (Dambis *et al.* 2001). As in the case of the Cepheid and OSC sample, the distance-scale correction factors given by the two methods differ systematically by $\sim$0.1. Table 6 lists the final kinematic parameters for the blue-sueprgiant sample.

Blue supergiants yielded somewhat higher angular velocity Ω_0 compared to what we inferred from the Cepheid and OSC sample, but the difference is within the quoted errors. The systematic difference between the two angular velocity values is partly due to the specifics of the space distribution of objects involved. The most reliable estimates of angular velocity are those inferred from objects lying in the vicinity of the

Table 8. Galactic rotation curve $V(R)$ for the short and long distance scales

R, kpc	$V(R)$, km s^{-1} $R_0 = 7.5$ kpc	$V(R)$, km s^{-1} $R_0 = 8.5$ kpc	R, kpc	$V(R)$, km s^{-1} $R_0 = 7.5$ kpc	$V(R)$, km s^{-1} $R_0 = 8.5$ kpc
2.0	198.1	198.2	8.2	201.6	227.5
2.1	199.9	201.3	8.3	200.7	227.0
2.3	201.3	203.8	8.4	199.8	226.4
2.4	202.4	205.9	8.6	199.0	225.8
2.6	203.3	207.6	8.7	198.3	225.2
2.7	204.0	209.0	8.9	197.6	224.6
2.8	204.6	210.3	9.0	196.9	224.0
3.0	205.2	211.4	9.1	196.3	223.3
3.1	205.7	212.3	9.3	195.8	222.7
3.3	206.2	213.2	9.4	195.4	222.1
3.4	206.7	213.9	9.6	195.1	221.5
3.5	207.2	214.7	9.7	194.8	220.9
3.7	207.8	215.4	9.8	194.6	220.4
3.8	208.3	216.2	10.0	194.5	219.9
4.0	208.9	216.9	10.1	194.5	219.4
4.1	209.5	217.6	10.3	194.6	219.0
4.2	210.1	218.4	10.4	194.8	218.7
4.4	210.6	219.1	10.5	195.0	218.5
4.5	211.2	219.9	10.7	195.4	218.3
4.7	211.7	220.7	10.8	195.8	218.2
4.8	212.2	221.5	11.0	196.3	218.1
4.9	212.6	222.3	11.1	196.9	218.2
5.1	212.9	223.1	11.2	197.6	218.3
5.2	213.2	223.8	11.4	198.4	218.5
5.4	213.4	224.6	11.5	199.2	218.9
5.5	213.5	225.3	11.7	200.2	219.3
5.6	213.5	226.0	11.8	201.2	219.8
5.8	213.4	226.6	11.9	202.4	220.4
5.9	213.3	227.2	12.1	203.6	221.1
6.1	213.0	227.7	12.2	205.0	221.9
6.2	212.7	228.2	12.4	206.4	222.8
6.3	212.2	228.6	12.5	208.0	223.7
6.5	211.7	229.0	12.6	209.8	224.8
6.6	211.1	229.2	12.8	211.6	226.0
6.8	210.5	229.4	12.9	213.7	227.2
6.9	209.7	229.5	13.1	215.9	228.5
7.0	208.9	229.6	13.2	218.3	229.9
7.2	208.1	229.6	13.3	220.9	231.4
7.3	207.2	229.5	13.5	223.7	233.0
7.5	206.3	229.3	13.6	226.7	234.6
7.6	205.4	229.0	13.8	230.0	236.3
7.7	204.4	228.8	13.9	233.6	238.0
7.9	203.5	228.4	14.0	237.4	239.8
8.0	202.5	228.0			

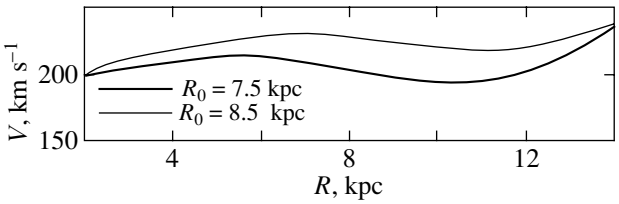

Fig. 2. Galactic rotation curve $V(R)$ for the short ($R_0 = 7.5$ kpc) and long ($R_0 = 8.5$ kpc) distance scales.

"tangent circle" (it is the circle in the Galactic plane with the interval connecting the Sun and the Galactic center as its diameter), because, in the corresponding conditional proper-motion equations, the coefficients at the angular-velocity derivatives are close to zero (Glushkova *et al.* 1999). The proper motions of 23 blue supergiants lying in the vicinity of the "tangent circle" yielded $\Omega_0 = (27.91 \pm 2.79)$ km s^{-1} kpc^{-1}; i.e., the angular velocities inferred from two samples agree well with each other. This result justifies the subsequent use of blue supergiants for constructing the combined rotation curve over a wide interval of Galactocentric distances.

Kinematics of Ionized Hydrogen

The only way to match the distance scales of H II and stars is to compare the first derivatives of angular velocity inferred from line-of-sight and space velocities for gas and stars, respectively. Given that the scatter of velocities along the z-coordinate has virtually no effect on the radial velocities of the thin-disk objects, we fixed $w_0 = -7$ km/s and $\sigma_w = \sigma_\nu$. Table 7 lists the kinematic and rotation-curve parameters inferred from H II data for $r <$ 4 kpc. In this interval of Galactocentric distances the first derivatives of angular velocity for gas (inferred from radial velocities) and stars (inferred from space velocities) are estimated at either -4.77 and -4.54 km s^{-1} kpc^{-2}, respectively (if $R_0 = 7.5$ kpc) implying the H II distance-scale correction factor $p =$ 0.95, or -4.08 and -3.66 km s^{-1} kpc^{-2}, respectively (if $R_0 = 8.5$ kpc) implying the H II distance-scale correction factor $p = 0.90$. Note that, as expected, the velocity ellipsoid axes inferred for gas do not obey the Lindblad relation, but $\sigma_u \approx \sigma_\nu$.

Constructing the Rotation Curve

The good agreement between the mean heliocentric velocity components of different young-object samples allows us to construct the rotation curve over a sufficiently wide interval of Galactocentric distances, 2–14 kpc, using radial velocities of both stars and gas. Figure 2 shows the rotation curves

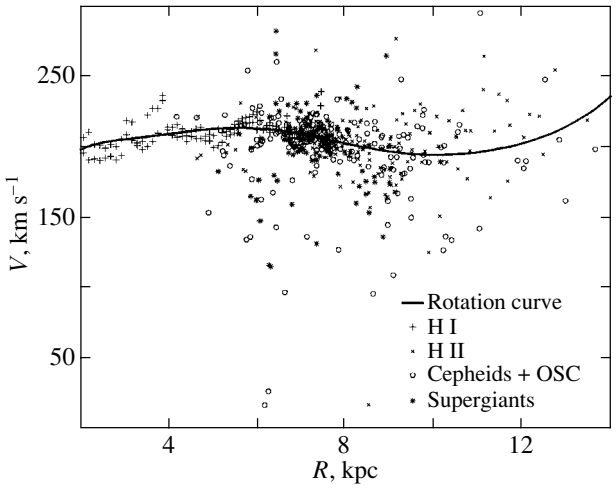

Fig. 3. Galactic rotation curve $V(R)$ for the short distance scale ($R_0 = 7.5$ kpc) with young-object data points superimposed.

$V(R)$ inferred from the entire sample of young objects (OSC + Cepheids + supergiants + H I + H II) for $R_0 = 7.5$ and 8.5 kpc. These rotation curves are tabulated in Table 8. Here we expanded the difference of angular velocities into a seventh- order Taylor series in the vicinity of R_0, and computed the velocity ellipsoid axes separately for neutral and ionized hydrogen, blue supergiants, and for Cepheids + OSC (see Table 9) with the distances to all objects matched to each other. The resulting local centroid velocity and Oort's constant A are equal to $V(R_0) = (206 \pm 10)$ km s^{-1}, $A = (17.1 \pm$ 0.5) km s^{-1} kpc^{-1} and $V(R_0) = (226 \pm 12)$ km s^{-1}, $A = (15.4 \pm 0.6)$ km s^{-1} kpc^{-1} for the short and long distance scale, respectively.

Figure 3 shows the rotation curve for the short distance scale ($R_0 = 7.5$ kpc) with the data points for individual objects computed using the following formula:

$$V = R\Omega_0 + \frac{R}{R_0 \sin l \cos b}(V_r - V_{\mathrm{sun},r}),$$

where $V_{\mathrm{sun},r}$ is the radial projection of the heliocentric velocity of the sample considered. As is evident from the figure, the scatter of V about the rotation curve is due mainly to small $\sin l$. Note that systematic differences between HI velocities may be manifestations of a barlike structure at the center of the Galaxy (Freudenreich 1998). The gas-stellar disk of the Galaxy is well known to show appreciable warp in the direction $l \approx 90°$ at Galactocentric distances >10 kpc. Our rotation curve therefore applies only to the part of the Galactic disk where warp is insignificant.

Table 9. Heliocentric velocity components and velocity-ellipsoid axes for the young-object sample

Objects	σ_u,	σ_ν,	σ_w^*,	u_0	ν_0	w_0^*
	km s^{-1}					
Cepheids + OSC	13.30	7.59	7.55	−9.17	−12.98	−7.25
Standard errors	±1.79	±0.41	−	»	»	»
Supergiants	14.17	10.00	5.13	»	»	»
Standard errors	±0.51	±2.12	−	»	»	»
H II	6.71	7.19	5.0	»	»	»
Standard errors	±0.60	±0.93	−	»	»	»
H I	6.60	6.05	5.0	»	»	»
Standard errors	−	±0.34	−	±0.48	±0.78	−

* Parameters fixed at values inferred from the space-velocity solution.

Table 10. Parameters of the spiral pattern ($R_0 = 7.5$ kpc)

Objects	m	f_R	f_θ	i	χ_0
		km s^{-1}		deg	
Cepheids	2	−6.66	−1.40	−6.02	−85.19
OSC	4	−5.51	−0.16	−12.18	−88.05
Standard errors		±2.34	±1.56	±0.72	±14.50
OB-stars	2	−6.64	0.42	−6.55	−97.28
Standard errors		±2.51	±2.31	±0.86	±18.30

Allowing for Spiral-Arm Effects

Our computations showed that young-object samples lag behind the Sun on the average by 13 km s^{-1}. Dehnen and Binney (1998) used HIPPARCOS proper motions and parallaxes of nearby MS stars to find out that the Sun moves ahead of the LSR by 5.25 km s^{-1}. It follows from this that young subsystems lag behind the LSR by $\sim$8 km s^{-1}, whereas their velocity dispersion should imply a velocity lag of $\leq$1.5 km s^{-1}. This discrepancy may be due, among other things, to streaming motions induced by spiral arms.

To allow for the spiral-arm effects in the velocity field, we performed our computations in terms of two- and four-armed models of the spiral pattern (see Table 10).

The phase of the Sun with respect to the spiral wave, which is close to $-\pi/2$, indicates that the Sun is situated at the outer edge of the arm; $V_\theta \approx$ 1.4 km s^{-1} (for the Cepheid + OSC sample). We therefore face a discrepancy between the magnitude and direction of the tangential disturbance as inferred from the centroid velocity lag behind the LSR ($V_\theta \approx$ −6.5 km s^{-1}) and the value of the same quantity computed directly in terms of a model of spiral-pattern effects in the velocity field. The discrepancy is beyond the quoted errors. The velocity lag of the centroid of young objects relative to the LSR can be explained by, among other things, noncircular motions of the LSR discussed by a number of authors (Schuter 1982; Clemens 1985) based on their analyses of H I radial velocities. Thus Clemens (1985) inferred an LSR tangential velocity of $\sim$7 km s^{-1} from an analysis of H I motions in the local solar neighborhood. However, Dehnen and Binney (1998) showed that, despite their different ages and velocity dispersions, all main-sequence stars (except late-type B-stars) closely follow a unified theoretical dependence of the sample tangential velocity on velocity dispersion. Streaming motions should be "washed out" by ever increasing velocity dispersion, and we therefore consider the above determinations of the solar velocity relative to the LSR to be quite correct, and thus the discrepancy in question is left unexplained. Interestingly, the heliocentric velocity of B-type stars inferred by Dehnen and Binney agrees well with the velocities we inferred for the young-object samples.

ACKNOWLEDGMENTS

We are grateful to A.M. Mel'nik, E.V. Glushkova, and V.G. Surdin for their assistance and valuable discussions. The work was supported by the Russian Foundation for Basic Research, grant nos. 01-02-06012, 00-02-17804, 99-02-17842, and 01-02-16086; Astronomy State Research and Technology Program; and the Council for the Support of Leading Scientific Schools, grant no. 00-15-96627.

REFERENCES

1. M. Barbier-Brossat and P. Figon, Astron. Astrophys., Suppl. Ser. **142**, 217 (2000).
2. H. Baumgardt, C. Dettbarn, and R. Wielen, Astron. Astrophys., Suppl. Ser. **146**, 251 (2000).
3. L. N. Berdnikov, O. V. Vozyakova, and A. K. Dambis, Pis'ma Astron. Zh. **22**, 936 (1996) [Astron. Lett. **22**, 838 (1996)].
4. R. Bottema, Astron. Astrophys. **275**, 16 (1993).
5. J. Brand and L. Blitz, Astron. Astrophys. **275**, 67 (1993).
6. D. P. Clemens, Astrophys. J. **295**, 422 (1985).
7. A. K. Dambis, Pis'ma Astron. Zh. **16**, 522 (1990) [Sov. Astron. Lett. **16**, 224 (1990)].
8. A. K. Dambis, Pis'ma Astron. Zh. **25**, 10 (1999) [Astron. Lett. **25**, 7 (1999)].
9. A. K. Dambis and A. S. Rastorguev, Pis'ma Astron. Zh. **27**, 132 (2001) [Astron. Lett. **27**, 108 (2001)].
10. A. K. Dambis, A. M. Mel'nik, and A. S. Rastorguev, Pis'ma Astron. Zh. **21**, 331 (1995) [Astron. Lett. **21**, 291 (1995)].
11. A. K. Dambis, A. M. Mel'nik, and A. S. Rastorguev, Pis'ma Astron. Zh. **27**, 68 (2001) [Astron. Lett. **27**, 58 (2001)].
12. A. K. Dambis, E. V. Glushkova, A. M. Mel'nik, and A. S. Rastorguev, Astron. Astrophys. Trans. **20**, 161 (2001).
13. W. Dehnen and J. J. Binney, Mon. Not. R. Astron. Soc. **298**, 387 (1998).
14. R. Drimmel and D. N. Spergel, Astrophys. J. **556**, 181 (2001).
15. M. Duflot, P. Figon, and N. Meyssonnier, Astron. Astrophys., Suppl. Ser. **114**, 269 (1995).
16. Yu. N. Efremov, *Sites of Star Formation in Galaxies: Star Complexes and Spiral Arms* (Nauka, Moscow, 1989).
17. M. Feast, F. Pont, and P. Whitelock, Mon. Not. R. Astron. Soc. **298**, L43 (1998).
18. J. Fernley, T. G. Barnes, I. Skillen, *et al.*, Astron. Astrophys. **330**, 515 (1998).
19. M. Fich, L. Blitz, and A. A. Stark, Astrophys. J. **342**, 272 (1989).
20. H. T. Freudenreich, Astrophys. J. **492**, 495 (1998).
21. E. V. Glushkova, A. K. Dambis, A. M. Mel'nik, and A. S. Rastorguev, Astron. Astrophys. **329**, 514 (1998).
22. E. V. Glushkova, A. K. Dambis, and A. S. Rastorguev, Astron. Astrophys. Trans. **18**, 349 (1999).
23. A. Gould and P. Popowski, Astrophys. J. **508**, 844 (1998).
24. S. L. Hawley, W. H. Jeffreys, T. G. Barnes III, and Lai Wan, Astrophys. J. **302**, 626 (1986).
25. M. Honma and Y. Sofue, Publ. Astron. Soc. Jpn. **49**, 453 (1997).
26. D. K. Karimova and E. D. Pavlovskaya, Astron. Zh. **50**, 737 (1973) [Sov. Astron. **17**, 470 (1973)].
27. S. Kent, T. M. Dame, and J. Fazio, Astrophys. J. **378**, 131 (1991).
28. P. N. Kholopov, Astron. Zh. **57**, 12 (1980) [Sov. Astron. **24**, 7 (1980)].
29. J. R. D. Lepine, Yu. N. Mishurov, and S. Yu. Dedikov, Astrophys. J. **546**, 234 (2001).
30. J. R. Lewis and K. C. Freeman, Astron. J. **97**, 139 (1989).
31. C. C. Lin, C. Yuan, and F. H. Shu, Astrophys. J. **155**, 721 (1969).
32. S. Malhotra, Astrophys. J. **448**, 138 (1995).
33. F. Maeder and G. Meynet, Astron. Astrophys., Suppl. Ser. **89**, 451 (1991).
34. A. M. Mel'nik, A. K. Dambis, and A. S. Rastorguev, Pis'ma Astron. Zh. **27**, 611 (2001) [Astron. Lett. **27**, 521 (2001)].
35. M. R. Merrifield, Astron. J. **103**, 1552 (1992).
36. Yu. N. Mishurov, E. D. Pavlovskaya, and A. A. Suchkov, Astron. Zh. **56**, 268 (1979) [Sov. Astron. **23**, 147 (1979)].
37. Yu. N. Mishurov, I. A. Zenina, A. K. Dambis, *et al.*, Astron. Astrophys. **323**, 775 (1997).
38. C. A. Murray, *Vectorial Astrometry* (A. Hilger, Bristol, 1983; Naukova Dumka, Kiev, 1986).
39. I. I. Nikiforov, Astron. Zh. **76**, 403 (1999) [Astron. Rep. **43**, 345 (1999)].
40. I. I. Nikiforov and I. V. Petrovskaya, Astron. Zh. **71**, 725 (1994) [Astron. Rep. **38**, 642 (1994)].
41. F. Pont, M. Mayor, and G. Burki, Astron. Astrophys. **285**, 415 (1994).
42. P. Popowski and A. Gould, Astrophys. J. **506**, 259 (1998).
43. A. S. Rastorguev, Determination of Rotation Curve and Scale of Distance in Galaxy, http://www.astronet.ru:8101/db/msg/1172553.
44. A. S. Rastorguev, E. V. Glushkova, A. K. Dambis, and M. V. Zabolotskikh, Pis'ma Astron. Zh. **25**, 689 (1999) [Astron. Lett. **25**, 595 (1999)].
45. A. S. Rastorguev, E. V. Glushkova, M. V. Zabolotskikh, and H. Baumgardt, Astron. Astrophys. Trans. **20**, 103 (2001).
46. K. Rohlfs, *Lectures on Density Wave Theory* (Springer-Verlag, Berlin, 1977; Mir, Moscow, 1980).
47. W. L. H. Shuter, Mon. Not. R. Astron. Soc. **199**, 109 (1982).
48. J. Torra, D. Fernandez, F. Figueras, and F. Comeron, Astrophys. Space Sci. **272**, 109 (2000).
49. T. Tsujimoto, M. Miyamoto, and Y. Yoshii, Astrophys. J. Lett. **492**, L79 (1998).
50. P. C. van der Kruit and K. C. Freeman, Astrophys. J. **303**, 556 (1986).

Translated by A. Dambis

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 465–475. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 529–540.

Analytic Formulas for the Mass-Transfer Rate and the Evolution of a Close Binary System of Neutron (Degenerate) Stars

V. S. Imshennik* and D. V. Popov**

Institute for Theoretical and Experimental Physics,
ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia

Received January 30, 2002

Abstract—We derive approximate analytic relations between the mass-transfer rate in a close binary system described in terms of the Roche potential and its basic parameters, such as the total mass of the binary, the radius of its circular orbit, the mass of the mass-losing component, and the degree of its Roche lobe overfilling. Using simplifying assumptions (conservative mass transfer, a short relaxation time of matter on the mass-gaining component compared to the mass-transfer time scale, adiabaticity and quasi-stationarity of the mass flow through the Lagrangian point L_1) allows the evolution of a binary system of neutron (degenerate) stars to be described in terms of two ordinary differential equations. This makes it possible to qualitatively analyze the evolution process, which is useful in those cases where the evolution of a close binary system must be investigated in general terms, for example, in terms of the scenario for the transformation of the collapse of a rotating presupernova core into a supernova explosion proposed by Imshennik and Nadyozhin (1992) and Imshennik (1992). © *2002 MAIK "Nauka/Interperiodica".*

Key words: *stars—structure and evolution*

INTRODUCTION

Mass transfer in semidetached close binary systems (CBSs) has been investigated in some detail in sophisticated numerical evolutionary models mostly for stars with ongoing thermonuclear reactions in their cores (see, e.g., the monograph by Masevich and Tutukov (1988) and references therein). In contrast, when studying CBSs composed of relativistic, primarily neutron stars (NSs), apart from complex numerical simulations of their evolution using complete three-dimensional hydrodynamic equations within the framework of general relativity (Oohara and Nakamura 1999; Fond *et al.* 2000), the need for constructing simplified, semianalytic models designed to estimate basic properties of their evolution with a wide range of possible binary parameters arises.

It should be emphasized that here, such models are valid for a sufficiently large mass difference between the CBS-forming NSs, and, in any case, they are inapplicable for equal masses of the CBS components. The need to construct such semianalytic models for the CBS evolution showed up most clearly when developing and investigating the scenario for the transformation of the collapse of a rotating presupernova iron core (including the collapse-induced fragmentation of this core to form a binary system of NSs and mass flow from the less massive component after the approach of the binary components) into its explosion with the explosion energies characteristic of collapsing supernovae (Imshennik and Nadyozhin 1992; Imshennik 1992, 1995).

The approach of the binary components and the ensuing mass transfer in a CBS composed of protoneutron (hot) stars have previously been studied in terms of this scenario by using simplified semianalytic models of mass transfer at the final stage, in particular, in the approximation of point like components when describing their gravitational interactions and radiation (for other simpler approximations, see below). These studies (Imshennik and Popov 1994, 1998) allowed us to find a satisfactory explanation for the temporal characteristics of the binary evolution and for the high initial pulsar escape velocities and to predict basic parameters of the gravitational radiation. These results were obtained with a wide range of binary parameters. The goal and content of our present study is to construct and justify such models for mass transfer in CBSs composed of degenerate NSs. Below, we give an analytic formula for the rate of mass transfer or, in other words, mass flow from one (less massive) component to the other (more massive) component. This formula was derived by using the well-known results from Paczynski and Sienkievich (1972), Savonier (1978), Pringle (1985),

*E-mail: Imshennik@vitep1.itep.ru
**E-mail: DVPopov@vitep1.itep.ru

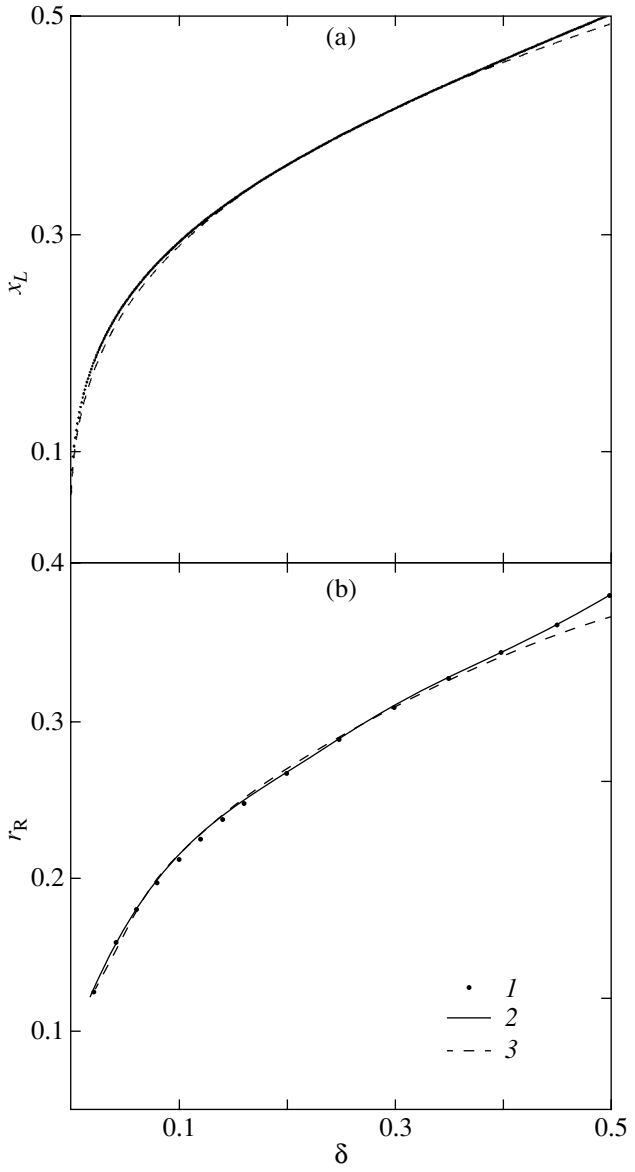

Fig. 1. Numerically determined dimensionless x_L (a) and Roche lobe radius (b) versus mass variable δ (*1*): *2* (b)—the Roche lobe radius calculated from Kopal's formula (5), *3*—the calculation from the approximate formulas (3) (a) and (4) (b).

Edwards and Pringle (1987), Ritter (1988), and Kolb and Ritter (1990), in which different and less general analytic relations for mass transfer in CBSs are presented. Next, we included the derived formula in a system of two ordinary differential equations that describes the evolution of a CBS of NSs with gravitational radiation and mass transfer almost until the mass of the less massive component decreases to a critical value of $\sim 0.1 M_\odot$.

A MODEL OF MASS TRANSFER IN A CBS. AN ANALYTIC EXPRESSION FOR THE MASS-TRANSFER RATE

The mass-transfer rate in a semidetached CBS was estimated (also analytically) by Paczynski and Sienkievich (1972), Savonier (1978), Pringle (1985), Edwards and Pringle (1987), Ritter (1988), and Kolb and Ritter (1990) with a gradual refinement of the model assumptions. The binary components in such models are assumed to be point masses, the orbit is assumed to be circular, and the Roche potential is introduced in the frame of reference rotating with the binary. This potential is the sum of the gravitational potentials of the point masses M_1 and M_2 and the centrifugal potential:

$$\Phi = -\frac{GM_1}{|\mathbf{r}-\mathbf{r_1}|} - \frac{GM_2}{|\mathbf{r}-\mathbf{r_2}|} - \frac{1}{2}\Omega^2(x^2+y^2).$$

Here, the coordinate system is defined in such a way that the orbital angular velocity Ω is directed along the z axis and the coordinate origin coincides with the center of mass. It is convenient to measure the lengths in units of the constant radius of a circular orbit a and the potential in units of GM_t/a, where the total mass of the binary components is denoted by $M_\mathrm{t} = M_1 + M_2$. Next, we choose one of the binary components, more specifically, the mass-losing component, as star 1 and call the second component star 2. Denote the relative mass of star 1 by $\delta = M_1/M_\mathrm{t}$. We place the coordinate origin somewhat differently, at the center of star 1, and direct the x axis toward star 2 located at point $(1,0,0)$. The center of mass of the binary then lies at point $(1-\delta,0,0)$. In this coordinate system, the expression for the dimensionless Roche potential is (Pringle 1985)

$$\Phi(x,y,z;\delta) = -\left[\frac{\delta}{\sqrt{x^2+y^2+z^2}} + \frac{(1-\delta)}{\sqrt{(1-x)^2+y^2+z^2}} + \frac{1}{2}[(x-1+\delta)^2+y^2]\right]. \quad (1)$$

To pass to dimensional units when calculating the derivatives of the potential below, we should multiply the first derivatives of Φ with respect to the coordinates by $\frac{GM_\mathrm{t}}{a^2}$ and the second derivatives by $\frac{GM_\mathrm{t}}{a^3} \equiv \Omega^2$, as suggested by the Kepler law.

The position of the first Lagrangian point (L_1) can be determined from the well-known condition $\left.\frac{\partial\Phi}{\partial x}\right|_{y=z=0} = 0$, $0 < x < 1$, which, after substituting the formula for Φ from (1), gives an equation for its coordinate $x_L(\delta)$:

$$\frac{\delta}{x_L^2} - \frac{1-\delta}{(1-x_L)^2} - (x_L - 1 + \delta) = 0. \quad (2)$$

Numerically solving this equation yields the value $x_L(\delta)$ used in our subsequent calculations, which for $\delta < 0.4$ can be satisfactorily (see Fig. 1) fitted by the formula

$$x_L(\delta) = 0.619\delta^{1/3}. \quad (3)$$

Note that in Fig. 1a, the results of our numerical solution merge into a solid line for $\delta > 0.05$. The following formula (Paczynski 1971) is commonly used to estimate the Roche lobe radius (also in units of a) defined as the radius of a sphere with volume equal to the Roche lobe volume over a wide range of mass ratios, $0.05 < \delta < 0.4$:

$$r_{\rm R}(\delta) = 0.462\delta^{1/3} \quad (4)$$

(the dimensional Roche lobe radius is $R_{\rm R} \equiv r_{\rm R}a$). The constancy of the ratio (equal to 0.764) of $r_{\rm R}$ to x_L calculated from formulas (3) and (4) implies that the Roche lobe shape is virtually invariable in this δ range; i.e., only its size varies. To more accurately fit $r_{\rm R}$ over the entire range $0 < \delta < 1$, we use the expression derived by Kopal (1959) when fitting the numerically calculated Roche lobe sizes (cf., e.g., Pringle 1985)

$$r_{\rm R}(\delta) = 0.49\left(\frac{\delta}{1-\delta}\right)^{2/3} \quad (5)$$
$$\times\left\{0.6\frac{\delta^{2/3}}{(1-\delta)^{2/3}} + \ln\left[1+\frac{\delta^{1/3}}{(1-\delta)^{1/3}}\right]\right\}^{-1}.$$

The numerically determined Roche lobe radius (in units of a) calculated by the integration of the Roche lobe volume and by the subsequent recalculation to the radius of a sphere with the corresponding volume is shown in Fig. 1. Also shown in the figure are the results yielded by the fits (4) and (5). We see that the fit of Kopal (1959) gives an excellent accuracy over the entire range of mass ratios δ under consideration, while the fit of Paczynski (1971) gives a good accuracy for $\delta < 0.4$.

The matter from star 1 that overfilled its Roche lobe will flow onto star 2. To determine the rate of this flow for a given degree of Roche lobe overfilling, we assume the following:

(1) Adiabaticity of the mass flow with a velocity $\mathbf{v}$ through the Lagrangian point L_1 and, accordingly, a polytropic equation of state for the matter in the outer envelope of the mass-losing star with a polytropic index n:

$$p = K\rho^{1+1/n}, \quad \varepsilon = np/\rho,$$

where K is the entropic constant.

(2) Quasi-stationarity of the flow, which can be checked by comparing the dynamical time scale in the outer envelope (characterizing the relaxation time of the stellar radius through mass loss) and the mass-transfer time scale.

(3) Potentiality of the mass flow in the vicinity of L_1; i.e., the absence of nonzero curl$\mathbf{v}$ here.

(4) The coincidence of streamlines with equipotential surfaces.

The latter assumption is partly confirmed by numerical simulations (Lubow and Shu 1975) of the gas flow near L_1. The deviation of streamlines from equipotential surfaces may be primarily due to the presence of Coriolis forces, which we ignore when introducing the Roche potential (1). The effect of these forces causes, in particular, the jet of flowing gas in the numerical calculations of Lubow and Shu (1975) to be inclined with respect to the line connecting the stellar centers at an angle that varies between 19° and 28°, depending on the binary parameters; in our model, however, this jet is directed along the x axis. Nevertheless, in estimating the mass-transfer rate, ignoring these effects seems admissible to a first approximation.

The above assumptions (adiabaticity, quasi-stationarity, potentiality of the flow in the vicinity of L_1, and the coincidence of streamlines with equipotentials) are used as the basis for our subsequent derivation. Under these assumptions, the Bernoulli equation is valid in the entire flow (Landau and Lifshitz 1986):

$$\frac{v^2}{2} + K(n+1)\rho^{1/n} = \Phi_{\rm s} - \Phi, \quad (6)$$

where Φ is the potential (1) and $\Phi_{\rm s}$ is its value on the surface of star 1, where the density becomes zero. Indeed, according to our assumptions, the streamlines on which the velocity becomes zero, because there is clearly a common point with a zero velocity on them (from symmetry considerations, it lies on the x axis on the other side of star 1 relative to L_1) passes over this surface. The potential $\Phi_{\rm s}$ is determined below. Note also that the absence of other integration constants for each streamline of the flow under consideration in Eq. (6) follows from the assumed flow potentiality (Landau and Lifshitz 1986).

Without changing the directions of the coordinate axes (the x axis is along the line connecting the stellar centers, the y axis is perpendicular to the rotation axis, and the z axis is along the rotation axis), for convenience of solving the problem, we move the coordinate origin to the first Lagrangian point L_1. We determine the density of the mass flux crossing the $x = 0$ plane, integrate it over the entire area crossed by the flow (it is specified by the condition $\Phi \le \Phi_{\rm s}$ under our assumptions), and, thus, find the total mass flux F identically equal to $-\dot{M}_1$.

Clearly, the streamlines will cross the $x = 0$ plane perpendicular to it (because they coincide with

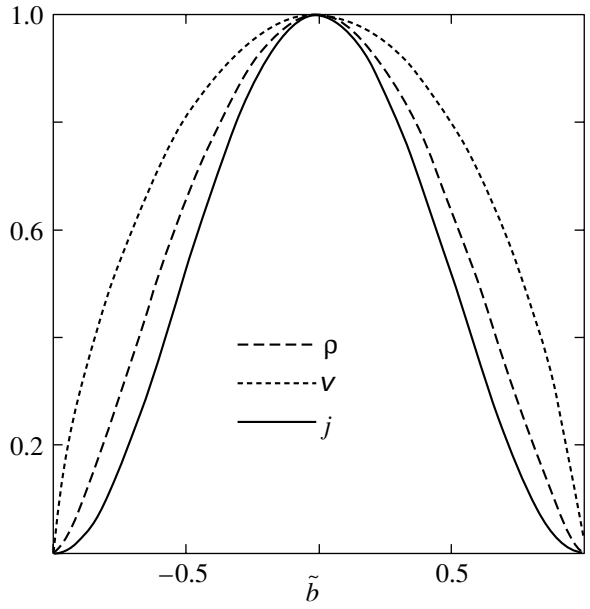

Fig. 2. The velocity (v), density (ρ), and mass flux density (j) profiles through the section at the Lagrangian point passing through a plane perpendicular to the binary axis for $n = 3/2$. Here, $b = \dfrac{y}{a_y b_s} = \dfrac{z}{a_z b_s}$.

equipotential surfaces). Note that the only equipotential passing through the point L_1 itself at an acute angle to the $x = 0$ plane is an exception, but its total contribution to the result obtained below is zero. Differentiating Eq. (6) along an arbitrary streamline and keeping in mind that the potential is constant on it, we obtain

$$v\frac{dv}{dx} + K\frac{n+1}{n}\rho^{-1+1/n}\frac{d\rho}{dx} = 0. \tag{7}$$

Since the streamlines beyond L_1 also coincide with equipotential surfaces, the flux per unit area on the $x = 0$ plane is at a maximum. This leads to an additional condition at $x = 0$:

$$\rho\frac{dv}{dx} + v\frac{d\rho}{dx} = 0. \tag{8}$$

Solving Eqs. (6)–(8) for the nonzero derivatives from (7) and (8) yields the velocity, density, and mass flux density $j = \rho v$ in the $x = 0$ plane:

$$\rho = \left[\frac{n(\Phi_s - \Phi)}{(n+1/2)K(n+1)}\right]^n, \tag{9}$$

$$v = \left(\frac{\Phi_s - \Phi}{n+1/2}\right)^{1/2}, \tag{10}$$

$$j = \left(\frac{\Phi_s - \Phi}{n+1/2}\right)^{(n+1/2)}\left[\frac{n}{K(n+1)}\right]^n. \tag{11}$$

As was shown, for example, by Landau and Lifshitz (1986) for the steady-state Euler equation, the condition (8) for flux density being at a maximum implies that the flow velocity is equal to the local speed of sound at a given point of the streamline, i.e., in our case, on the entire $x = 0$ plane. Indeed, the velocity v given by Eq. (10) is identically equal to the speed of sound $C = \left(K\dfrac{n+1}{n}\rho^{1/n}\right)^{1/2}$ for the equation of state under consideration. The numerical calculations of Lubow and Shu (1975) revealed that the sonic point (i.e., the point at which the flow velocity is equal to the local speed of sound) is very close to the Lagrangian point L_1. This fact suggests that the above assumptions are physically justifiable. For example, the flow nonadiabaticity and nonpotentiality attributable to the presence of shock fronts must certainly take place, but all these fronts are located in the supersonic flow region, beyond L_1. Therefore, an approximate analytic calculation of the mass-transfer rate can be carried out under our assumptions, which greatly simplifies the problem.

Let us now calculate the positive difference between the dimensional potentials, $\Phi_s - \Phi$, by assuming this difference to be small. Let the potentials be measured from L_1. Expanding the potential near L_1 in a Taylor series and realizing that its first derivatives with respect to all coordinates are zero at the Lagrangian point, we find that in the $x = 0$ plane,

$$\Phi(y,z) = \frac{GM_t}{a}[h_y(\delta)y^2 + h_z(\delta)z^2], \tag{12}$$

where y and z are measured in units of a and the second derivatives were calculated from Eq. (1):

$$h_y(\delta) = \frac{1}{2}\frac{\partial^2\Phi}{\partial y^2}\bigg|_{L_1} \tag{13}$$
$$= \frac{1}{2}\left[\frac{\delta}{x_L(\delta)^3} + \frac{1-\delta}{(1-x_L(\delta))^3} - 1\right],$$

$$h_z(\delta) = \frac{1}{2}\frac{\partial^2\Phi}{\partial z^2}\bigg|_{L_1} \tag{14}$$
$$= \frac{1}{2}\left[\frac{\delta}{x_L(\delta)^3} + \frac{1-\delta}{(1-x_L(\delta))^3}\right].$$

Here, it should be noted that the right-hand sides of Eqs. (13) and (14) contain the quantity $x_L(\delta)$ determined from Eq. (2), which can be fitted by (3) for $\delta < 0.4$.

Thus, according to Eq. (12), the equipotential lines in the $x = 0$ plane are similar ellipses

$$\frac{y^2}{a_y^2} + \frac{z^2}{a_z^2} = b^2,$$

where $a_y = 1/\sqrt{h_y}$, $a_z = 1/\sqrt{h_z}$, and the similarity coefficient is denoted by b. Hence, the sought quantity $\Phi_s - \Phi$ is, as can be easily verified (see (12)), a function of one variable, $b_s^2 - b^2$:

$$\Phi_s - \Phi = \frac{GM_t}{a}(b_s^2 - b^2), \tag{15}$$

where b_s is ultimately determined by the degree of Roche lobe overfilling. Substituting (15) in (11) yields j as a function of b:

$$j(b) = \left(\frac{GM_t}{a}\right)^{(n+1/2)} \times \frac{n^n}{K^n(n+1)^n(n+1/2)^{(n+1/2)}}(b_s^2 - b^2)^{n+1/2}.$$

This function normalized to its value at the jet center, i.e., at $b = 0$, is plotted in Fig. 2 for the typical $n = 3/2$. Similarly, substituting (15) in (9) and (10) yields ρ and v as functions of b. These functions, also normalized to their values at $b = 0$, are shown in Fig. 2 (for $n = 3/2$). Integrating $j(b)$ over the area of the flow section in the $x = 0$ plane under consideration and using the fact that the differential of the ellipse area: $dS = a^2\pi a_y a_z 2b\ db$ (we insert the dimensionless factor a^2 in dS) depends on b, we obtain for the total mass flux

$$F = \int j\ dS = a^2 \int_0^{b_s} j(b)\pi a_y a_z 2b\ db$$
$$= \left(\frac{GM_t}{a}\right)^{n+1/2} \frac{n^n}{K^n(n+1)^n(n+1/2)^{n+1/2}} \times \frac{2\pi a^2}{\sqrt{h_y(\delta)h_z(\delta)}} b_s^{2n+3} \int_0^1 (1-\xi^2)^{n+1/2}\xi\ d\xi,$$

whence we easily find

$$F = \left(\frac{GM_t}{a}\right)^{n+1/2} \times \frac{n^n}{K^n(n+1)^n(n+1/2)^{n+1/2}} \times \frac{2\pi a^2}{\sqrt{h_y(\delta)h_z(\delta)}} \frac{b_s^{2n+3}}{2n+3}. \tag{16}$$

It is more convenient to rewrite this equation by substituting b_s derived from Eq (15) with $\Phi = 0$ at $b = 0$:

$$F = \frac{2\pi a^3}{GM_t} \times \frac{n^n}{K^n(n+1)^n(n+1/2)^{(n+1/2)}(2n+3)} \times \frac{\Phi_s^{n+3/2}}{\sqrt{h_y(\delta)h_z(\delta)}}. \tag{17}$$

Thus, the sought mass-transfer rate (the mass flux from the mass-losing component) is expressed in terms of the excess of the surface potential above the potential of L_1. The sunsequent calculations yield expressions for this flux that include only the mass of star 1 and model parameters. Therefore, the excess of the potential should be related to the mass of star 1. Note that for the Roche potential, there is a unique relationship between the potential on some equipotential surface and the volume within it. We assume that there is a unique relationship between radius (in other words, volume) and mass for a single, spherically symmetric star. Let us now assume that when such a star of a given mass moves into the Roche potential (i.e., into a binary system), its volume will not change (a refinement of this assumption requires three-dimensional hydrodynamic calculations of the stellar structure in the binary system similar to those of Kuznetsov (1995) and is not made here). If the mass of star 1 corresponds to a volume smaller than the Roche lobe volume (at a given separation a between the components), then no mass transfer takes place; if, alternatively, it corresponds to a volume slightly larger than the Roche lobe volume, then the Roche lobe is overfilled and mass transfer takes place. As long as there is no mass transfer (the star volume is smaller than the Roche lobe volume), there is a unique relationship between the volume occupied by the star and the potential on its surface. Thus, we can numerically determine the derivative of the potential on the surface of such a star with respect to its volume. When the surface potential is larger than the potential at L_1, the volume within the equipotential surface undergoes an abrupt change (the Roche lobe of star 2 is added). However, it is physically clear that at low degrees of Roche lobe overfilling, star 1 as a whole does not feel this transition, and to determine the relationship between the potential on its surface and the degree of Roche lobe overfilling, the relationship between surface potential and volume should be extended without a jump in the derivative from the range of volume corresponding to Roche lobe underfilling to the volumes corresponding to its slight overfilling.

The sought derivative of the surface potential with respect to the volume within the surface can be numerically determined by computing a series of models close to Roche lobe filling. The inferred left-hand derivative at $\Phi = \Phi(L_1)$ should then be continuously extended to potentials slightly larger than $\Phi(L_1)$. Thus, we find that

$$\Phi_s = \left.\frac{d\Phi}{dV}\right|_{L_1} \Delta V.$$

This formula allows the mass flux (17) to be related to the Roche lobe overfilling volume ΔV. The derivative

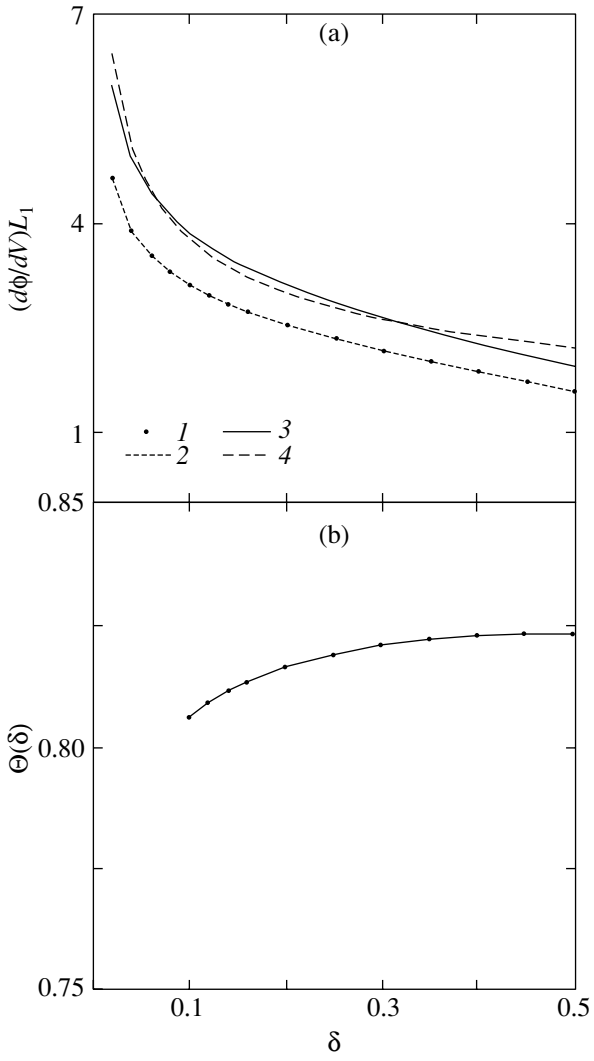

Fig. 3. (a) The numerically determined (for various δ) derivatives of the Roche potential on an equipotential surface with respect to the volume within this surface at the potential of L_1 (*1*); *2*—interpolation; the derivative was calculated on the left, i.e., at a lower potential (and volume); the units of measurement GM_t/a^4; *3*—the derivative calculated in the spherically symmetric approximation (18), with the Roche lobe volume being determined from Kopal's formula (5); *4*—a similar calculation but the Roche lobe volume was calculated from Paczynski's formula (4). (b) The ratio $\Theta(\delta)$ of this numerically determined derivative to the derivative calculated from Kopal's formula (5). The equation of state for the matter, including the polytropic index n, has no effect on these plots.

of the surface potential with respect to the volume within the surface used in the formula is shown in Fig. 3a. It is important to note that although the Roche lobe radius is determined from formula (5) (Kopal 1959) with excellent accuracy, when substituting the corresponding derivative for a spherically symmetric potential

$$\frac{d\Phi}{dV} = \frac{1}{3}\left(\frac{4\pi}{3}\right)^{1/3}\frac{GM_t\delta}{V^{4/3}} \qquad (18)$$

at the volume equal to the Roche lobe volume calculated from formula (5), the introduced error is $\sim 20\%$ (Fig. 3b). We may conclude that this uncertainty in the excess of the sought derivative compared to its more accurate three-dimensional determination shown in Fig. 3a is quantitatively not too large. It is also important that the ratio of the derivatives of the potential with respect to the volume (found by an accurate numerical calculation and determined from the spherization procedure) is essentially constant: when δ changes over the entire range of interest, i.e., from 0.1 to 0.5, this ratio changes from 0.806 to 0.823 (see Fig. 3b). Denoting this ratio by $\Theta(\delta)$, we obtain (in dimensional units)

$$\Phi_s \approx \Theta(\delta)\frac{GM_1(R_1 - R_R)}{R_R^2} \qquad (19)$$
$$\approx \Theta(\delta)\frac{GM_1(R_1 - R_R)}{R_1 R_R}.$$

We emphasize once again that in deriving Eq. (19), we assumed a low degree of Roche lobe overfilling and a continuity of the derivative of the surface potential with respect to the volume within the surface when its values are extended in the case of Roche lobe overfilling under consideration. It should be noted that the one-dimensional derivative $\left.\frac{d\Phi}{dx}\right|_{L_1}$ is zero at the Lagrangian point L_1 by definition and cannot be used to determine the potential Φ_s. On the other hand, there would be little sense in further increasing the accuracy of expression (19) by solving a self-consistent problem of the configuration of star 1 in the Roche potential. Moreover, it would become an excess of the accuracy relative to the use of the Roche lobe radius in the previous derivation based on the spherization of the volume occupied by star 1. Of course, there is the problem of the assumption of a continuous derivative $\left.\frac{d\Phi}{dV}\right|_{L_1}$ being forced, because we can say nothing certain about the star configuration when the Roche lobe is appreciably overfilled.

Substituting the approximate value of Φ_s from (19) in (17) yields an expression for the total mass flux $F \equiv -\dot{M}_1$,

$$-\dot{M}_1 = \frac{a^3}{GM_t} \qquad (20)$$
$$\times \frac{n^n}{K^n(n+1)^n(n+1/2)^{n+1/2}(2n+3)}$$

$$\times \frac{2\pi}{\sqrt{h_y(\delta)h_z(\delta)}}$$
$$\times \left(\frac{GM_1\Theta(\delta)}{R_1}\right)^{n+3/2}\left(\frac{R_1-R_\mathrm{R}}{R_\mathrm{R}}\right)^{n+3/2}.$$

Next, let us consider simpler models used to estimate $\dot{M}_1$ by Pringle (1985) and Edwards and Pringle (1987) and compare the above results with the estimates of these models. These models ignore the macroscopic flow of gas through L_1 and assume that the gas particles diffuse into a vacuum with the thermal velocity equal to the local speed of sound. The density and the speed of sound at L_1 are estimated under the same assumptions as above [see Eq. (6) and its justification] but with $v = 0$. In other words, for the matter at L_1, the parameters are assumed to be the same as those for a point with the same potential far from L_1.

Using Eq. (6) with $v = 0$, we find that the density and the speed of sound at L_1 ($\Phi = 0$) are related to the potential Φ_s of the stellar surface relative to L_1:

$$\rho = \frac{\Phi_\mathrm{s}^n}{K^n(n+1)^n}, \quad C = \sqrt{\frac{\Phi_\mathrm{s}}{n}}.$$

Next, following the simplified model under consideration, we assume that the flux near L_1 also has an elliptical cross section determined by the condition that its specific internal ($C^2/2$) and potential ($\Phi_\mathrm{s}^{(P)}$) energies being equal at its boundary; hence, the potential of the flow surface in our estimate is related to the previously introduced one by

$$\Phi_\mathrm{s}^{(P)} = \frac{\Phi_\mathrm{s}}{2n}.$$

Since $\Phi_\mathrm{s} \propto b_\mathrm{s}^2$, i.e., the potential is proportional to the flow area, the area in this approach is underestimated by a factor of $2n$ (by a factor of 3 for $n = 3/2$). Given that the flux per unit area in the estimate of Pringle (1985) is constant, we find, after integration over the area, that the total mass flux given by the simplified model differs from that calculated above (see (20)) by a factor that depends on n alone:

$$\kappa(n) = \frac{(2n+3)(n+1/2)^{n+1/2}}{4n^{n+3/2}}. \qquad (21)$$

This quantity is shown in Fig. 4a; it is $\kappa(3/2) = 16/9$ for the important case of $n = 3/2$. Given the roughness of the models under consideration, this difference cannot be fundamental, although we see that the simplified model of Pringle (1985) yields an overestimated mass-transfer rate for any n. Formally, for absolutely rigid matter, $n = 0$ ($\gamma = \infty$), the factor $\kappa(n)$ goes to infinity, yielding an infinite mass-transfer rate in the simplified estimate. This is because the speed of sound becomes infinite, giving a physically meaningless infinite mass flux. A more careful analysis leading to expression (20) does not have this drawback, because $\lim\limits_{n\to 0} n^n = 1 < \infty$. However, for physically realistic adiabatic indices $\gamma = 1 + 1/n$ of the stellar matter, the two approaches convey the dependences of the total mass flux on basic physical binary parameters (the component masses, the binary size, and the degree of Roche lobe overfilling) in virtually the same way.

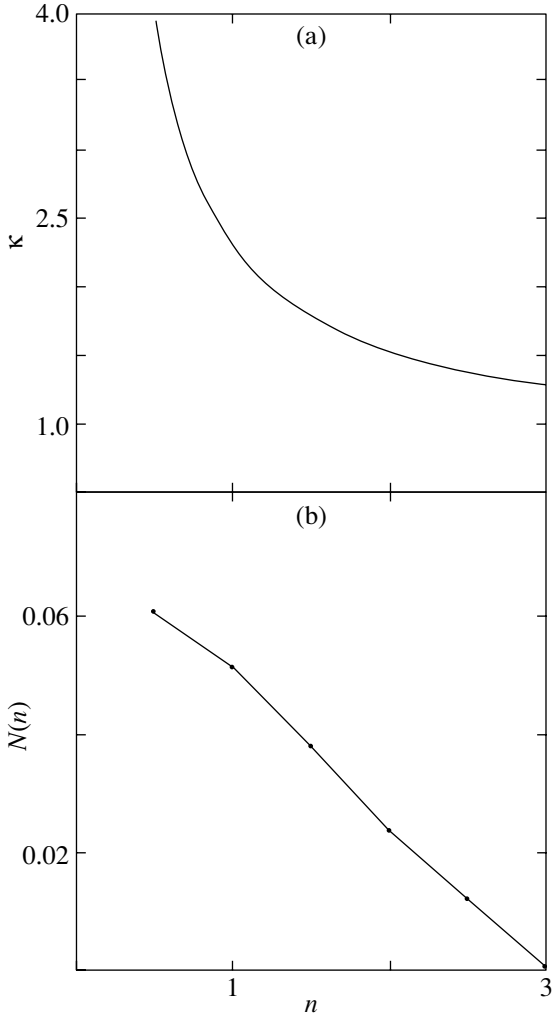

Fig. 4. (a) The ratio of the mass flux of flowing matter in the simplified model of Pringle (1985) to its more accurate estimate (20) [$\kappa(n)$ introduced in (21)] and (b) the function $N(n)$ introduced in (24) versus the polytropic index of flowing matter n.

To derive the final equation that could be used in semianalytic models of the CBS evolution, let us express the entropic constant K in terms of the radius R_1 and mass M_1 of star 1. This can be done if the entire star is assumed to be a polytrope with a single

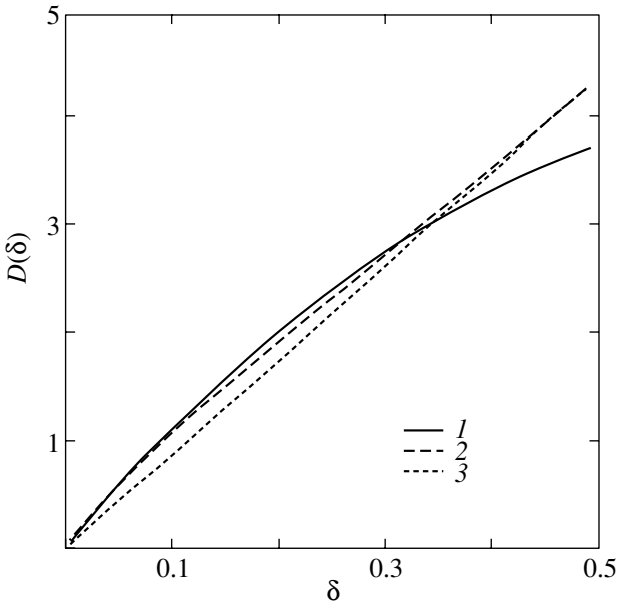

Fig. 5. The plot of $D(\delta)$; Kopal's more accurate fit (5) (*1*) and Paczynski's more approximate formula (4) (*2*) compared with its fit by a linear function $D = 8.5\ \delta$ (*3*) were used to calculate $r_{\rm R}(\delta)$.

adiabatic index.[1] Then, we have from the standard formulas for polytropic models

$$K^n = G^n M_1^{n-1} \frac{R_1^{3-n}}{\tilde{R}(n)^{3-n}},$$

where $\tilde{R}(n)$ is the form-factor of the polytrope that depends on its index alone, which can be found in standard tables (see, e.g., Chandrasekhar 1939). We assume that star 1 is stable against collapse, i.e., $n < 3$. Substituting K^n in (20), we write the relation between the total mass flux in the model under consideration and the parameters of the mass-losing star 1 as

$$-\dot{M}_1 = 2\pi G^{1/2} \times \frac{n^n \tilde{R}(n)^{3-n}}{(n+1)^n (n+1/2)^{n+1/2}(2n+3)} \frac{a^3}{M_{\rm t}} \times \frac{\Theta(\delta)^{n+3/2}}{\sqrt{h_y(\delta)h_z(\delta)}} \frac{M_1^{5/2}}{R_1^{9/2}} \left(\frac{R_1 - R_{\rm R}}{R_{\rm R}}\right)^{n+3/2}. \quad (22)$$

Dividing the two sides of Eq. (22) by the constant (in the conservative model of mass transfer under consideration) total binary mass and using a specified expression for the Roche lobe radius ((4) or (5)), we finally rewrite formula (22) in a form more convenient for our calculations, namely, as a differential equation for the variable quantity δ:

$$\dot{\delta} = -2\pi G^{1/2} M_{\rm t}^{1/2} N(n) D(\delta) a^{-3/2} \left(\frac{R_{\rm R}}{R_1}\right)^{9/2} \times \left(\frac{R_1 - R_{\rm R}}{R_{\rm R}}\right)^{n+3/2}. \quad (23)$$

Here, we use the possibility (illustrated in Fig. 3b) of substituting the above constant for $\Theta(\delta) \approx 0.815$ with an accuracy sufficient for our analysis and denote the auxiliary functions by

$$N(n) = \frac{n^n \tilde{R}(n)^{3-n} 0.815^{n+3/2}}{(n+1)^n (n+1/2)^{n+1/2}(2n+3)} = \frac{0.815^{n+3/2}}{4\kappa(n)} \frac{\tilde{R}(n)^{3-n}}{(n+1)^n n^{3/2}}, \quad (24)$$

$$D(\delta) = \frac{\delta^{5/2}}{\sqrt{h_y(\delta)h_z(\delta)} r_{\rm R}(\delta)^{9/2}}. \quad (25)$$

The function $N(n)$ is plotted in Fig. 4b. The plot of $D(\delta)$ obtained after substituting the exact (in the entire range $0 < \delta < 0.5$) fit (5) in (25) is shown in Fig. 5. This plot leads us to conclude that $D(\delta)$ can be easily fitted with a linear function $D(\delta) \approx 8.5\delta$, which is suitable for qualitatively estimating the behavior of the solution to the equations of the mass-transfer model under consideration. This is because $\Theta(\delta) \approx 0.815 \pm 0.01$ is virtually constant with δ; h_y and h_z are only slightly varying functions of δ, and $r_{\rm R}(\delta) \propto \delta^{1/3}$ for moderately large values of the argument [see (4)].

USING THE DERIVED EXPRESSION FOR THE MASS-TRANSFER RATE TO ANALYTICALLY MODEL THE CBS EVOLUTION

Let us calculate the evolution of the radius of a binary orbit. Assuming that the total binary mass is conserved and that the angular momentum is lost through the radiation of gravitational waves alone (which is of importance for CBSs composed of relativistic stars, for example, (proto)neutron stars), we obtain the following equation by using the standard treatment of the Newtonian dynamics of the motion of a mass point with a variable mass (see, e.g., Jaranowski and Krolak 1992):

$$\frac{da}{dt} = -\frac{64G^3 M_{\rm t} M_1 M_2}{5c^5 a^3} - 2a\frac{(M_2 - M_1)}{M_1 M_2}\frac{dM_1}{dt}. \quad (26)$$

[1] Note, however, that we need not make this strong assumption. In contrast to the surface layers of star 1 with a polytropic index n, we could introduce a different polytropic index $n^* \neq n$ for the entire star, which is implied in the ensuing relations for the entropic constant K.

Here, the first term on the right-hand side describes the radiation of gravitational waves in the approximation of two point masses (Peters and Mathews 1963; Landau and Lifshitz 1973), which carries energy and angular momentum away from the binary. The second term on the right-hand side of Eq. (26) reflects the effect of mass transfer from star 1 to star 2 on the rate of change in orbital radius. To describe this mass transfer, we use a conservative model (Clark and Eardley 1977) in (26) with the total mass and total angular momentum of the binary conserved. As was shown by Blinnikov *et al.* (1984), this model is applicable if the time of angular-momentum transfer from the accretion disk to star 2 is shorter than the mass-transfer time scale and if, in addition, the mass and angular momentum of the matter that leaves the binary through the outer Lagrangian point L_2 may be neglected. It is easy to verify that without allowance for the effect of gravitational radiation, conservative mass transfer from the less massive component always increases the total energy of the binary $E = -GM_1M_2/2a$. This increase can be produced by the internal energy of hot star 1.

The system of model equations can be simplified when the radius of star 1 is a function of its mass; the degree of its Roche lobe overfilling is then determined by its mass and by the binary radius. Therefore, the problem can be reduced to solving two ordinary differential equations for the relative mass δ of star 1 and the degree of its Roche lobe overfilling, i.e., the ratio of its radius to the Roche lobe radius. In this case, the binary radius can be expressed in terms of these quantities, and the temporal evolution of the CBS can be easily investigated qualitatively. It is of interest to carry out such a study of a CBS composed of relativistic stars, for example, in terms of the scenario for the collapse of a rotating presupernova core (Imshennik and Nadyozhin 1992; Imshennik 1992, 1995), which, through fragmentation, leads to the formation of a CBS of protoneutron stars, and to mass loss by the less massive component down to the lower limit of NS masses and its explosion as a collapsing supernova. Our previous study of this process (Imshennik and Popov 1998) is based on the equations for the model of mass transfer in a CBS derived in this paper.

The radius of cool NS 1 (R_1) is determined by its mass alone; thus, it may be assumed to be a function of δ, which, of course, also depends on the parameter $m_\mathrm{t} \equiv M_\mathrm{t}/M_\odot$ constant in each calculation. We use the following formula to calculate R_1:

$$R_1 = R_0 \frac{m_1^{\beta_1}}{(m_1 - \mu)^{\beta_2}} \qquad (27)$$

(the fit to the results of the calculations by Baym *et al.* (1971) and Jaranowski and Krolak (1992)), where the basic numerical values of the parameters are $R_0 = 7.5$ km, $\beta_1 = 0.79$, $\beta_2 = 0.83$, $\mu = 0.09$, and $m_1 \equiv M_1/M_\odot$. Note that, first, the relation between the NS mass and radius in form (27) was derived for the equation of state for cool NSs; for a more faithful description of "loose" hot or rotating protoneutron stars with much lower densities and larger radii, we must increase R_0 possibly to $\sim$20–30 km (we emphasize the need to perform careful calculations of the protoneutron-star cooling within the first hours after collapse). Second, this fit correctly conveys the fact that the NS radius becomes infinite when its mass decreases to the lower limit of NS masses μ, approximately equal to $0.09M_\odot$. Qualitatively, this is equivalent to its explosion, obtained, for example, for a sufficiently rapid decrease in mass in the model of Blinnikov *et al.* (1990).[2] Of course, the fact that the radius R_1 becomes infinite at $m_1 = \mu$ has no quantitative physical meaning, because this radius for a NS with a very small excess of its mass above the critical value is about 200 km (Blinnikov *et al.* 1990; Aksenov *et al.* 1995). Therefore, the "truncation" of formula (27) must be implied at $R_1 \approx 200$ km.

Let us determine the function $\eta(\delta) \equiv R_1/(R_0 r_\mathrm{R}(\delta))$, where R_1 and R_0 are introduced in (27) and $r_\mathrm{R}(\delta)$ is specified by formula (5) (or the simplified formula (4)); the quantity m_t appears in the function η as a parameter. When using expression (27) to calculate the NS radius and to determine $r_\mathrm{R}(\delta)$ according to (4), we have:

$$\eta(\delta) = \frac{\delta^{\beta_1 - 1/3}}{0.462\, m_\mathrm{t}^{\beta_2 - \beta_1} (\delta - \mu/m_\mathrm{t})^{\beta_2}}. \qquad (28)$$

If expression (5) is used to calculate r_R, then the formula for $\eta(\delta)$ appears much more complex, but the results for $\delta \le 0.4$ are virtually the same. Consequently, we can perform numerical calculations using formula (5) and a qualitative analysis of our main results using (4).

Denote the quantity that characterizes the degree of Roche lobe overfilling by $\xi = R_\mathrm{R}/R_1$. Using the expression derived from the definition of ξ and $\eta(\delta)$ for

$$a = R_0\, \xi\, \eta(\delta) \qquad (29)$$

and taking into account the definition of $r_\mathrm{R}(\delta)$, we rewrite Eq. (23) in the new variables ξ and δ:

$$\frac{d\delta}{dt} = -\frac{2\pi N(n)\sqrt{GM_\odot}\, m_\mathrm{t}^{1/2}}{R_0^{3/2}} \qquad (30)$$

[2] The exact critical mass of low-mass NSs appears to be known with an accuracy of no less than 3%; it depends on the equation of state for neutron matter, on the kinetics of β processes, and even on general-relativity effects. Its more accurate value can be slightly higher than the value taken in the text: $0.095M_\odot$ (Blinnikov *et al.* 1990; Aksenov *et al.* 1995).

$$\times \frac{D(\delta)}{\eta(\delta)^{3/2}}\,\xi^3\,\left(\frac{1-\xi}{\xi}\right)^{n+3/2}.$$

In the special but physically interesting case $n = 3/2$, Eq. (30) is

$$\frac{d\delta}{dt} = -\frac{2\pi N(3/2)\sqrt{GM_\odot}m_t^{1/2}}{R_0^{3/2}} \times \frac{D(\delta)}{\eta(\delta)^{3/2}}\,(1-\xi)^3. \quad (31)$$

Let us now express M_1 and M_2 in Eq. (26) in terms of δ and M_t, substitute a from (29) and the derivative of δ from (30), and, as a result, rewrite Eq. (26) as

$$\frac{d\xi}{dt} = -\frac{64G^3M_\odot^3m_t^3}{5c^5}\frac{\delta(1-\delta)}{R_0^4\eta(\delta)^4}\frac{1}{\xi^3} + \xi\left[\frac{2(1-2\delta)}{\delta(1-\delta)} + \frac{d\ln\eta(\delta)}{d\delta}\right]\frac{2\pi N(n)\sqrt{GM_\odot}m_t^{1/2}}{R_0^{3/2}} \times \frac{D(\delta)}{\eta(\delta)^{3/2}}\,\xi^3\,\left(\frac{1-\xi}{\xi}\right)^{n+3/2}, \quad (32)$$

where the logarithmic derivative can be explicitly expressed from (28) for the fit (4) of the Roche lobe radius:

$$\frac{d\ln\eta(\delta)}{d\delta} = \frac{\beta_1 - 1/3}{\delta} - \frac{\beta_2}{(\delta - \frac{\mu}{m_t})}; \quad (33)$$

for the more accurate fit (5), this derivative is substituted in (32) numerically.

For an arbitrary polytropic index, Eqs. (30) and (32) describe the evolution of a CBS of (proto)neutron stars with gravitational radiation and mass transfer when the mass fraction of the less massive component $\delta = \delta(t)$ changes from an initial value at $t = 0$ $\delta_0 < 0.5$ to the critical value $\delta = \mu/m_t$ at some time $t = t_f$. Formally, Eqs. (30) and (32) form a system of two differential equations for two unknown functions $\delta(t)$ and $\xi(t)$ with the initial conditions $\xi = 1$ (mass transfer begins when the radius of star 1 is equal to the Roche lobe radius) and $\delta = \delta_0$ (δ_0 can take on any value between μ/m_t and 0.5; in principle, it must be determined by calculating the fragmentation of a collapsing rotating presupernova core). The right-hand sides of these equations depend in a complicated way both on the function δ—on which, in turn, depend the given functions $\eta(\delta)$, $d\ln\eta(\delta)/d\delta$, $D(\delta)$—as well as on the function ξ; in contrast to δ, this latter dependence is explicit only. The given constant quantities R_0 and M_t (it is more convenient to use the dimensionless total mass m_t), as well as the world constants G and c, appear on the right-hand sides as dimensional parameters. In addition, $N(n)$ (a dimensionless function of the polytropic index n which appears in the second term of Eq. (32)) enters into the right-hand side of Eq. (30). It may be noted that in the important special case of $n = 3/2$, it is $N(3/2) = 0.0378$.[3]

CONCLUSIONS

The dynamics of a binary star system is known to be completely determined by the specification of three parameters, for example, the total mass of the binary (M_t), the orbital radius (a), and the component-mass ratio ($\delta/(1-\delta)$), provided that the components are in a circular orbit (in the Newtonian approximation of gravitational interaction for point masses). If the total mass of the binary is conserved (here, we considered conservative mass transfer), then two parameters are required to determine the binary dynamics. Therefore, it is easy to understand why describing the CBS evolution reduces to solving the two differential equations (30) and (32) that were derived under certain simplifying assumptions described in detail above. In this case, it should certainly be borne in mind that the radius R_1 of the mass-losing component is assumed to be uniquely related to its mass. Indeed, mass transfer includes an additional evolutionary parameter, the degree of Roche lobe overfilling $\xi = \xi(t)$: Apart from the mass fraction of the less massive component $\delta = \delta(t)$, the latter characterizes the binary evolution and can be expressed, by definition, in terms of the radius R_1.

The derived system of ordinary differential equations (30) and (32) allows us to perform qualitative and quantitative analysis of the CBS evolution, which was previously carried out for a binary system of protoneutron stars (Imshennik and Popov 1998). It can also be used to model other CBSs (primarily those

[3]When using a similar model in our previous paper (Imshennik and Popov 1998), where we numerically solved the above system of Eqs. (30) and (32) and applied our results to qualitatively model the evolution of a CBS of protoneutron stars, we used a simplified procedure for determining the relationship between Φ_s and the degree of Roche lobe overfilling, which essentially implied setting $\Theta(\delta)$ equal to unity. As a result, $N(n)$ was calculated with an error approximately equal to $0.815^{n+3/2}$; for the important special case of $n = 3/2$, $N(n)$ was overestimated by a factor of 1.85. However, our calculations with a refined $N(n)$ confirm our main conclusions (Imshennik and Popov 1998). This is because such a refinement decreases the mass-transfer rate for a given degree of Roche lobe overfilling, for example, by only a factor of 2 for $n = 3/2$, leaving the model evolution close to the limiting case of finite mass transfer for infinitesimal Roche lobe overfilling considered previously (Imshennik and Popov 1996). In this case, it proved to be possible to analytically obtain the major predicted observational manifestations of the evolution of a CBS composed of (proto)neutron stars. This suggestion can be easily checked by numerically solving the system of Eqs. (30) and (32).

composed of relativistic stars). This analysis makes it possible to qualitatively estimate the effect of basic physical binary parameters on the CBS evolution and can be used in combination with more detailed numerical simulations of the evolution.

ACKNOWLEDGMENTS

We wish to thank A.G. Aksenov, S.I. Blinnikov, and D.K. Nadyozhin for helpful discussions. We are grateful to R.A. Sunyaev and K.A. Postnov for useful critical remarks. D.V. Popov thanks N.A. Bobrova for great interest in the results and for support.

REFERENCES

1. A. G. Aksenov, S. I. Blinnikov, and V. S. Imshennik, Astron. Zh. **72**, 717 (1995) [Astron. Rep. **39**, 638 (1995)].
2. G. Baym, H. A. Bethe, and C. J. Petchick, Nucl. Phys. A **175**, 225 (1971).
3. S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, and A. G. Polnarev, Pis'ma Astron. Zh. **10**, 422 (1984) [Sov. Astron. Lett. **10**, 177 (1984)].
4. S. I. Blinnikov, V. S. Imshennik, D. K. Nadezhin, *et al.*, Astron. Zh. **67**, 1181 (1990) [Sov. Astron. **34**, 595 (1990)].
5. S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (Univ. of Chicago Press, Chicago, 1939).
6. J. P. A. Clark and D. M. Eardley, Astrophys. J. **215**, 311 (1977).
7. D. A. Edwards and J. E. Pringle, Mon. Not. R. Astron. Soc. **229**, 383 (1987).
8. J. A. Fond, M. Miller, W.-M. Suen, and M. Tobias, Phys. Rev. D **61**, 044011 (2000).
9. V. S. Imshennik, Pis'ma Astron. Zh. **18**, 489 (1992) [Sov. Astron. Lett. **18**, 194 (1992)].
10. V. S. Imshennik, Space Sci. Rev. **74**, 325 (1995).
11. V. S. Imshennik and D. K. Nadyozhin, Pis'ma Astron. Zh. **18**, 195 (1992) [Sov. Astron. Lett. **18**, 79 (1992)].
12. V. S. Imshennik and D. V. Popov, Pis'ma Astron. Zh. **20**, 620 (1994) [Astron. Lett. **20**, 529 (1994)].
13. V. S. Imshennik and D. V. Popov, Preprint No. 940, MPA (1996).
14. V. S. Imshennik and D. V. Popov, Pis'ma Astron. Zh. **24**, 251 (1998) [Astron. Lett. **24**, 206 (1998)].
15. P. Jaranowski and A. Krolak, Astrophys. J. **394**, 586 (1992).
16. U. Kolb and H. Ritter, Astron. Astrophys. **236**, 385 (1990).
17. Z. Kopal, *Close Binary Systems* (Wiley, New York, 1959).
18. O. A. Kuznetsov, Astron. Zh. **72**, 508 (1995) [Astron. Rep. **39**, 450 (1995)].
19. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics,* Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
20. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics,* Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
21. S. H. Lubow and F. H. Shu, Astrophys. J. **198**, 383 (1975).
22. A. G. Masevich and A. V. Tutukov, *Evolution of Stars: Theory and Observations* (Nauka, Moscow, 1988).
23. K. Oohara and T. Nakamura, Proc. Theor. Phys. Suppl. **136**, 270 (1999).
24. B. Paczynski, Annu. Rev. Astron. Astrophys. **9**, 183 (1971).
25. B. Paczynski and R. Sienkievich, Acta Astron. **22**, 73 (1972).
26. P. C. Peters and J. Mathews, Phys. Rev. **131**, 435 (1963).
27. J. E. Pringle, in *Interacting Binary Stars,* Ed. by J. E. Pringle and R. A. Wade (Cambridge Univ. Press, Cambridge, 1985), Part 1.
28. H. Ritter, Astron. Astrophys. **202**, 93 (1988).
29. G. W. Savonier, Astron. Astrophys. **62**, 317 (1978).

Translated by V. Astakhov

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 476–487. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 541–553.

On the Production of Chemical Elements Beyond the Iron Peak

I. V. Panov[1]* and V. M. Chechetkin[2]

[1] *Institute for Theoretical and Experimental Physics, ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia*

[2] *Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya pl. 4, Moscow, 125047 Russia*

Received January 30, 2002

Abstract—We discuss the possibility of elements heavier than iron being produced in the cooling central part of a low-mass supernova remnant in terms of a consistent kinetic model that incorporates reactions with neutrons, protons, α particles and β decay. We show that at typical density ($\sim 10^9$ g cm^{-3}) and temperature ($\sim 5 \times 10^9$), chemical elements with atomic masses up to $A \sim 130$ can be produced not only in the classical r-process but also through a consistent allowance for β decay in nuclear-equilibrium calculations in a medium with a large neutron excess. The chemical elements from iron to nuclides with masses $A \sim 130$ can be produced under these conditions, which makes up for a deficit in the yields of chemical elements in current models for the r-process. If the initial neutron excess is large ($\eta_0 \geq 0.4$), then the nuclear equilibrium will be disturbed during matter expansion, and a decrease in temperature and the dynamical process of rapid nucleosynthesis can take place. For this nucleosynthesis scenario, the production of chemical elements from iron to uranium is probable even at low initial neutron densities.

Key words: *nuclear astrophysics, nucleosynthesis; supernovae and supernova remnants*

INTRODUCTION

Heavy nuclei beyond the iron peak are known to be produced in nature mainly through neutron-capture reactions (Burbidge *et al.* 1957). Analysis of the Solar-system abundance curve (Cameron 1982) shows that elements heavier than iron are synthesized under the action of neutrons in two different processes characterized by different conditions. The first (s) process takes place when the rates of β decay are much higher than the rates of (n, γ) reactions: $\lambda_\beta \gg \lambda_{n\gamma}$ (for neutron densities $n_\mathrm{n} \sim 10^{16}$ cm^{-3}). The mechanism of the s-process is well understood (Käppeler *et al.* 1989). The second (r) process takes place under conditions of high neutron densities and temperatures such that $\lambda_\beta \ll \lambda_{n\gamma}$; the nuclei involved in such nucleosynthesis have a large neutron excess and a short lifetime. The numerous studies of this process carried out in the past 40 years clearly revealed the conditions required for heavy nuclei to be synthesized (see, e.g., Käppeler *et al.* (1998) and references therein). However, the detailed scenario for the r-process is not yet completely understood.

Over the period of studying the r-process, more than ten scenarios (see, e.g., Mathews and Cowan 1990) have been proposed for this process, including an explosion on the neutron-star surface (Bisnovatyi-Kogan and Chechetkin 1979), a collision of a neutron star with a black hole (Lattimer and Schramm 1976), and an explosion of a low-mass neutron star (Imshennik 1992). The conditions typical of the r-process can be achieved, in particular, during type II supernova explosions, for example, for high-mass stars in the model with an artificial piston (Hillebrandt 1978; Woosley and Hoffman 1992). The conditions required for the r-process are also achieved during neutron-star mergers (Lattimer and Schramm 1974; Symbalisty and Schramm 1982; Freiburghaus *et al.* 1999a). In this scenario, however, some of the physical parameters are also model-dependent.

As was mentioned above, the details of the mechanism for the formation of conditions for the r-process are not yet clear, but in recent years, substantial progress has been made in this problem. The yields of elements obtained in r-process calculations (Freiburghaus *et al.* 1999b) and observations of old stars (Sneden *et al.* 2000) show that there are at least two groups of physical scenarios for the r-process (Wasserburg *et al.* 1996; Qian *et al.* 1998). The first group (where the main r-process is realized) assumes the production of chemical elements with

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

atomic masses $A > 120$. The necessary conditions can be obtained both in the scenario of supernova-shell ejection by a hot neutrino wind (Woosley and Hoffman 1992; Takahashi *et al.* 1994) and during neutron-star mergers (Rosswog *et al.* 1999). The scenario for the production of elements heavier than iron with atomic masses $60 < A < 120$ (an additional r-process) is not yet clear. There are several models: the helium-flash model (Hillebrandt and Thielemann 1977; Truran *et al.* 1978), the model of rapid nucleosynthesis induced by a neutrino pulse from a collapsing supernova (Epstein *et al.* 1988; Nadyozhin *et al.* 1998), and the model for the production of r-elements in the central region of a low-mass type-I supernova with a large neutron excess $\eta = (N - Z)/A$ and, accordingly, with low $Y_e = (1 - \eta)/2$ (Hillebrandt *et al.* 1984; Ptitsyn and Chechetkin 1982; Wheeler *et al.* 1998) during the deflagration burning of a CO core discussed by Ptitsyn and Chechetkin (1982) and Panov *et al.* (1995).

In our view, elements with atomic masses $60 < A < 120$ are produced during an explosion of a low-mass supernova not only through (n, γ) reactions but also through other reactions with protons and α particles (at least during the initial stage). Therefore, we often use the term rapid nucleosynthesis rather than the r-process (Panov and Nadyozhin 1999). Under actual conditions of the initial stage of the scenario under consideration (at high temperatures and densities), nucleosynthesis proceeds mainly through charged-particle reactions. As the temperature and density decrease and as the charged-particle reactions decrease in importance, such nucleosynthesis gradually transforms into the dynamical r-process in the presence of a sufficient number of free neutrons. The production of some of the heavy nuclei ($Z > 26$) before the onset of the r-process significantly relaxes the requirements both on the neutron source (a smaller number of free neutrons is required) and on the seed nuclei (they are formed in charged-particle reactions). In the scenarios for the r-process developed in recent years, rapid nucleosynthesis begins immediately after the formation of seed nuclei in the α process (Woosley and Hoffman 1992) during an explosion of a high-mass supernova or during neutron-star mergers (Rosswog *et al.* 1999). In the above papers, the α- and r-processes are modeled separately, in terms of various mathematical models and codes, which imposes certain constraints on the physical models in the range of solution-joining parameters. Thus, in particular, when modeling the r-process in terms of a unified nucleosynthesis model, the problem of a correct allowance for the absorption of neutrons by light elements—"poisons" with a large neutron capture cross section—can arise.

Since the new scenarios (see above) for rapid nucleosynthesis suggest a significant role of charged particles in producing heavy elements, such nucleosynthesis cannot be reliably analyzed in the waiting point approximation or the approximation of nuclear statistical equilibrium. A consistent investigation of rapid nucleosynthesis requires a full reaction network that, apart from neutron reactions, incorporates charged-particle reactions and that allows one to determine the effect of charged particles on rapid nucleosynthesis under any of the conditions obtained in evolutionary models in those cases where this effect is significant and to determine these conditions. In addition, by analyzing rapid nucleosynthesis in terms of a single model over a wide temperature range and with various nuclear reaction channels being automatically switched on and off, we can avoid simplifying the physics of the phenomenon when solving the problem under consideration.

THE PRODUCTION OF CHEMICAL ELEMENTS FROM IRON TO XENON

Model r-process calculations satisfactorily reproduce the observed heavy-element abundance curve (Käppeler *et al.* 1998; Witti *et al.* 1993; Woosley *et al.* 1994; Freiburghaus *et al.* 1999b; Blinnikov and Panov 1996). However, in nucleosynthesis calculations for specific scenarios, such as a supernova explosion or a neutron-star merger, there is virtually no yield of heavy elements with $A < 120$ (Freiburghaus *et al.* 1999a). Therefore, of interest are the r-process models in which the nucleosynthesis conditions are not sufficient for the main r-process to proceed but which are suitable for the additional r-process responsible for the synthesis of elements with $70 < A < 120$. Without purporting to comprehensively review such models, we briefly consider three likely scenarios for the additional r-process.

(1) The helium-flash model proposed a quarter of a century ago (Hillebrandt and Thielemann 1977; Truran *et al.* 1978) assumes that an incomplete (additional) r-process takes place when a shock wave passes through the helium shell of a supernova with a mass of $\sim 10 M_\odot$. Through the heating of the medium behind the shock front, neutrons are intensively produced in the $^{18}\mathrm{O}(\alpha, n)^{21}\mathrm{Ne}$ and $^{22}\mathrm{Ne}(\alpha, n)^{25}\mathrm{Mg}$ reactions and the r-process can begin on the seed nuclei formed in the s-process and in equilibrium nucleosynthesis. It was assumed (because no detailed calculations were performed) that, depending on the details of the scenario, the yield of r-elements could be in good agreement with observations.

(2) For supernovae with masses $M > 8\mathrm{M}_\odot$, whose explosions are accompanied by core collapse with the

emission of an intense neutrino impulse, rapid nucleosynthesis in the supernova shell can be triggered by the neutrinos produced during the collapse of the internal stellar shells. In particular, as was shown by Epstein *et al.* (1988), the neutrino scattering by helium nuclei followed by their breakup can give rise to the source of the neutrons required for the r-process.

This model was considered in detail by Nadyozhin *et al.* (1998) and Panov and Nadyozhin (1999). They showed that the r-process in low-metallicity stars with masses $M \leq 15M_\odot$ could produce chemical elements with atomic masses $80 < A < 120$.

(3) Yet another schematic model (Ptitsyn and Chechetkin 1982), when worked out in detail, could give a realistic scenario for the additional r-process. Here, it should be noted that the above authors were perhaps the first to pay attention (although incompletely) to the role of protons in producing seed nuclei in the r-process. The explosion model of a type I supernova with a CO core (Ivanova *et al.* 1977; Chechetkin *et al.* 1980) forms the basis for the scenario under consideration. In this model, through the total breakup of the star, the matter of its central region with a high neutron excess ($Y_e \leq 0.3$) is ejected into outer space. In this neutronized matter, the conditions needed for the r-process to begin and proceed can be achieved.

As calculations show, the conditions created during explosions of such low-mass stars are suitable for the production of elements between the first and second peaks on the abundance curve through charged-particle reactions and, in several cases, are also sufficient for the production of heavier elements up to $A \sim 196$. At the onset of explosion, much of the matter consists of iron-peak nuclei in equilibrium. As the layers of a low-mass star expand (see Ivanova *et al.* (1983) for details) and as the matter cools to below a temperature of $\sim 6 \times 10^9$ K, the equilibrium conditions are violated and the conditions for the synthesis of elements heavier than iron with $A > 80$ can be realized in the presence of a sufficient number of neutrons.

To estimate these conditions, Panov *et al.* (1995) and Blinnikov *et al.* (1995) developed a model of the r-process that incorporates the (γ, p) and (p, γ) reactions in which the conditions necessary for the the r-process to begin are formed in the ejected shells of type Ia supernovae. When the temperature is high enough ($T > 2 \times 10^9$ K), the (p, n) and (n, p) reactions and the reactions involving α particles cannot be ignored. Therefore, the results of Panov *et al.* (1995) and Blinnikov *et al.* (1995), who, like Ptitsyn and Chechetkin (1982), took into account the (p, γ) and (γ, p) reactions alone in addition to reactions with neutrons and β decay, were only preliminary: as was shown recently (Panov and Nadyozhin 1999), the role of reactions with α particles can also be prominent, particularly if there are a sufficient number of free α particles, for example, in the helium shell.

A self-consistent solution of the nucleosynthesis problem under astrophysical conditions is known to require that the scenarios for the r-process leading to the production of heavy elements in the reactions of rapid neutron capture by nuclei correspond to stellar evolutionary models that determine the conditions for the r-process to proceed. In particular, one of the weak points in the available models for the production of heavy elements is the absence of conditions under which a high density of neutrons and the density of seed nuclei required for the r-process are reached simultaneously. Usually, in the models under consideration, either the seed nuclei are too few in number and are rapidly depleted or the number of neutrons per seed nucleus is insufficient for the synthesis of heavy elements. Thus, apart from the conditions necessary for the r-process to proceed (temperatures, densities, neutron number densities), it is necessary to find the production sources of neutrons and seed nuclei.

A possible solution for the above problem, at least for nuclei near the first and second peaks on the abundance curve, was pointed out by Ptitsyn and Chechetkin (1982). For the possible initial conditions of the r-process they suggested conditions close to the equilibrium ones but with an enhanced degree of matter neutronization (the r-process under these conditions was called the *rbc*-process). Under the physical conditions typical of the *rbc*-process, two groups of nuclei can exist (Ptitsyn and Chechetkin 1982). The first group with $Z < Z^*$ (where $\lambda_{\gamma p} \sim \lambda_{p\gamma}$ and $\lambda_{\gamma n} \sim \lambda_{n\gamma}$) is the most abundant group of nuclei being produced in the e-process and forming a maximum of the equilibrium distribution. For the other, less abundant group of nuclei with $Z > Z^*$ (where $\lambda_{\gamma p} \sim \lambda_\beta$), the condition of detailed balance is no longer satisfied for the reactions with neutrons, protons, and α-particles (at $\lambda_\beta \geq \lambda_{n\gamma}$). The equilibrium number densities $n(A, Z)$ break down, resulting in the formation of new nuclei with larger Z:

$$(A + 1, Z) \rightarrow (A + 1, Z + 1) + e^- + \tilde{\nu}.$$

Panov *et al.* (1995) formulated the problem in more detail, and Blinnikov *et al.* (1995) began to investigate this problem in terms of a consistent kinetic model developed to solve nucleosynthesis problems (Blinnikov and Panov 1996). However, these authors did not take into account the many reactions with α-particles that must play a significant role in the scenario for the explosion of a star and for the production of new nuclei under consideration either.

Here, we bridge the gap for reactions with protons and α-particles, study the dependence of the results on the neutronization of the initial medium, and determine the mechanism for the production of chemical elements with $80 < A < 130$ for specific conditions ($T \sim 5 \times 10^9$, $\rho \sim 10^9$ g cm^{-3}).

The present calculations differ from our previous calculations, in particular, by more complete and accurate nuclear data. More specifically, a common approach combining experimental data (where possible) with consistent theoretical reaction-rate calculations (see Cowan *et al.* (1991) for pair reactions and Kratz *et al.* (1993) for β decay) was used to calculate the neutron cross sections (for detailed references on nuclear data, see, e.g., Panov *et al.* (2001b)). The principal difference between the present and previous calculations is that the yield of elements with $A > 130$ is much smaller than that in the calculations of Panov *et al.* (1995) and Blinnikov *et al.* (1995). It is determined mainly by allowance for all principal reactions and by new nuclear data, in particular, by a more realistic mass formula (Hilf *et al.* 1976), which gives a considerably smaller number of existing isotopes for elements with $Z > 50$.

In addition, we analyze the edge effect that was disregarded previously (Panov *et al.* 1995; Blinnikov *et al.* 1995). In the above papers, because of an incomplete data bank for the charged-particle reaction rates (only for $Z < 45$), the (p, X) reactions produced a leak of matter from the equilibrium range ($Z < 45$) into the range where only the reactions with neutrons and β decay ($Z > 45$) were taken into account.

THE MODEL AND THE SYSTEM OF EQUATIONS

Thus, we consider the conditions that arise during the explosion and breakup of a low-mass supernova with a CO core. As was shown by Imshennik *et al.* (1999), no detonation takes place in the degenerate matter in the central region. Therefore, the burning proceeds in deflagration regime, which provides a high degree of matter neutronization and a low entropy before the total breakup of the stellar core.

At the onset of explosion, much of the matter consists of iron-peak nuclei in equilibrium with the following typical initial temperatures and densities: $T_9^0 \approx 5$ ($T_9 \equiv T/10^9$ K) and $\rho_0 \approx 1 \times 10^9$ g cm^{-3}. As the stellar layers expand and cool down below a temperature of (5–6) $\times 10^9$ K, the equilibrium conditions are violated, the effect of β decay shows up, and the conditions for the synthesis of elements heavier than iron with $A > 80$ are created.

As the initial time ($t = 0$) we choose the time that corresponds to the expansion and cooling of the supernova remnants; since our prime objective was to elucidate the behavior and role of β decay in the synthesis of elements for the conditions under consideration, the detailed astrophysical scenario is unimportant. For simplicity, we disregarded the temporal variations in temperature and density in most of our calculations and considered only the effect of a neutron excess in the initial matter on the mass distribution of the forming nuclei.

In our calculations, we used a previously developed (Blinnikov and Panov 1996; Nadyozhin *et al.* 1998) kinetic model of nucleosynthesis in which the change in the fraction $Y_{A,Z}$ of each nuclide (A, Z) was described by the differential equation

$$\begin{aligned} dY_{A,Z}/dt &= -\lambda_\beta(A,Z)Y_{A,Z} \\ &- \lambda_{\mathrm{n}\gamma}(A,Z)Y_{A,Z} + \lambda_{\mathrm{n}\gamma}(A-1,Z)Y_{A-1,Z} \\ &- \lambda_{\gamma\mathrm{n}}(A,Z)Y_{A,Z} + \lambda_{\gamma\mathrm{n}}(A+1,Z)Y_{A+1,Z} \\ &- \lambda_{\mathrm{p}\gamma}(A,Z)Y_{A,Z} + \lambda_{\gamma\mathrm{p}}(A+1,Z+1)Y_{A+1,Z+1} \\ &- \lambda_{\gamma\mathrm{p}}(A,Z)Y_{A,Z} + \lambda_{\mathrm{p}\gamma}(A-1,Z-1)Y_{A-1,Z-1} \\ &- \lambda_{\alpha\gamma}(A,Z)Y_{A,Z} + \lambda_{\gamma\alpha}(A+4,Z+2)Y_{A+4,Z+2} \\ &- \lambda_{\gamma\alpha}(A,Z)Y_{A,Z} + \lambda_{\alpha\gamma}(A-4,Z-2)Y_{A-4,Z-2} \\ &- \lambda_{np}(A,Z)Y_{A,Z} + \lambda_{np}(A,Z+1)Y_{A,Z+1} \\ &- \lambda_{\mathrm{pn}}(A,Z)Y_{A,Z} + \lambda_{\mathrm{pn}}(A,Z-1)Y_{A,Z-1} \\ &- \lambda_{\mathrm{p}\alpha}(A,Z)Y_{A,Z} + \lambda_{\alpha\mathrm{p}}(A-3,Z-1)Y_{A-3,Z-1} \\ &- \lambda_{\alpha\mathrm{p}}(A,Z)Y_{A,Z} + \lambda_{\mathrm{p}\alpha}(A+3,Z+1)Y_{A+3,Z+1} \\ &+ \lambda_{\alpha\mathrm{n}}(A-3,Z-2)Y_{A-3,Z-2} - \lambda_{\alpha\mathrm{n}}(A,Z)Y_{A,Z} \\ &- \lambda_{\mathrm{n}\alpha}(A,Z)Y_{A,Z} + \lambda_{\mathrm{n}\alpha}(A+3,Z+2)Y_{A+3,Z+2} \\ &+ \lambda_{\nu e}(A,Z-1)Y_{A,Z-1} - \lambda_{\nu e}(A,Z)Y_{A,Z} \\ &+ \sum_{k=0,1,2,3} \lambda_\beta(A+k,Z-1) \\ &\times P_k(A+k,Z-1)Y_{A+k,Z-1}, \end{aligned} \tag{1}$$

where λ specify the rates of various processes. For β decay,

$$\lambda_\beta(A,Z) = \ln 2/T_{1/2}(A,Z).$$

For two-particle (i, j) reactions,

$$\lambda_{ij}(A,Z) = Y_i(t)\rho(t)N_A\langle\sigma_{ij}(A,Z)v\rangle,$$

where $i = n, p, \alpha$ and $j = n, p, \alpha, \gamma$.

The rates of reverse photodissociation reactions $\lambda_{\gamma i}$ were determined via the rates of direct (i, γ) reactions (Fowler and Hoyle 1967):

$$\lambda_{\gamma i}(A,Z) = C_{AZ}^i N_A\langle\sigma_{i\gamma}(A-A_i, Z-Z_i)v\rangle,$$

where

$$C_{AZ}^i = 0.987 \times 10^{10}\frac{g_i g_{A-Ai,Z-Z_i}}{g_{A,Z}}$$

$$\times \left[\frac{A_i(A-A_i)}{A}T_9\right]^{3/2} \exp\left(-Q_{i\gamma}\frac{11.605}{T_9}\right)\ \mathrm{g\,cm^{-3}},$$

$A_i = 1, 1, 4$, $Z_i = 0, 1, 2$, and $g_i = 2, 2, 1$ for n, p and α, respectively, and $Q_{i\gamma}$ is the energy (in MeV) released in the corresponding direct (n,γ)-, (p,γ)-, and (α,γ)-reactions: $(A-A_i, Z-Z_i) + i \rightarrow (A,Z) + \gamma + Q_{i\gamma}$.

The fractions of neutrons, protons, and α particles are described by the following kinetic equations, which must be solved simultaneously with system (1):

$$dY_n/dt = -\sum_Z\sum_A[\lambda_{n\gamma} - \lambda_{\gamma n} + \lambda_{np} - \lambda_{pn} \quad (2)$$
$$+ \lambda_{n\alpha} - \lambda_{\alpha n} - \sum_{k=1,2,3} k\lambda_\beta(A,Z)P_k(A,Z)]Y_{A,Z},$$
$$dY_p/dt = -\sum_Z\sum_A(\lambda_{p\gamma} - \lambda_{\gamma p} + \lambda_{p\alpha}$$
$$- \lambda_{\alpha p} + \lambda_{pn} - \lambda_{np})Y_{A,Z},$$
$$dY_\alpha/dt = -\sum_Z\sum_A(\lambda_{\alpha\gamma} - \lambda_{\gamma\alpha} + \lambda_{\alpha p} + \lambda_{\alpha n}$$
$$- \lambda_{n\alpha} - \lambda_{p\alpha})Y_{A,Z},$$

where

$$Y_{A,Z} = \frac{n(A,Z)}{\rho N_A}, \quad Y_n = \frac{n_n}{\rho N_A}, \quad (3)$$
$$Y_p = \frac{n_p}{\rho N_A}, \quad Y_\alpha = \frac{n_\alpha}{\rho N_A}.$$

Compared to the equations presented by Blinnikov and Panov (1996) and Nadezhin *et al.* (1998), the system of Eqs. (1) and (2) contain the additional terms that describe the reactions involving α particles (as in the paper by Panov and Nadyozhin (1999)).

In contrast to the calculations of Nadezhin *et al.* (1998), we take into account the emission of one, two, and three delayed neutrons in Eq. (1): $P_0 + P_1 + P_2 + P_3 = 1$, where P_0 is the probability of β decay without any emission of delayed neutrons, and P_1, P_2, P_3 are the emission probabilities of one, two, and three neutrons after β decay, respectively.

The total number of nuclides and, hence, of Eqs. (1) for nucleosynthesis calculations generally depends on the choice of boundary conditions; here, the number is ~ 3000. We specified the boundary conditions for the range of nuclei in which nucleosynthesis was modeled as follows: $Z_{\min} = 6$ (carbon) and $Z_{\max} = 82$ (lead); the minimum and maximum A for each Z were specified by the boundaries of neutron and proton stability, which were determined from the mass formula. The nuclear masses used were obtained by Kratz *et al.* (1993) with allowance for Coulomb corrections and with a more accurate allowance for pairing effects (Möller and Nix 1992) compared to other mass formulas. The range with the magic number of neutrons $N = 82$, where the masses from Hilf *et al.* (1976) were used, constituted an exception (for more details, see Kratz *et al.* (1993)).

The (n,γ)-reaction rates, as well as the photodissociation reaction rates, were taken mainly from Thielemann *et al.* (1987) and Cowan *et al.* (1991). The mass relations and β-decay rates (when no experimental data were available), as with the emission probabilities of delayed neutrons, were determined from predictions of the QRPA model described in detail by Möller *et al.* (1997). For $Z < 46$, the rates of reactions with protons and α particles were calculated by using $N_A\langle\sigma v\rangle$ taken from the data library of Thielemann *et al.* (1987) and supplemented for $Z > 46$ with the reaction rates calculated in terms of the same formalism by Rauscher and Thielemann (2000).

Note that of the weak processes, we considered only β decay. Although the temperature is significant, β decay was considered only from the ground state of atomic nuclei.

The electron and positron captures that cause the neutron excess to change were not considered either. It is clear from available estimates of the rates of these processes (Martinez-Pinedo *et al.* 2000) that the reaction of electron capture by a proton has the maximum reaction rate among the processes listed above. In our case, however, the proton density is low, and on actual nucleosynthesis time scales (0.1–1 s), although the results of our calculations may be sensitive to allowance for this process, this process will not cause any appreciable change in the results (and conclusions).

The pair production cross section at $T_9 = 5$ is small; accordingly, the positron density is much lower than the electron density (Fowler and Hoyle 1967). The positron-capture reaction rate is $\sim 10^{-3}$ s^{-1} (Lang 1978), i.e., it is small and appreciably less than the rate of β decay. In addition, since the number of free protons is two to three orders of magnitude smaller than that of neutrons, the increase in the number of neutrons is negligible, even at an appreciable positron density. However, if this increase took place, an increase in the density of free neutrons could only enhance the production of heavier elements as the degree of initial neutronization of the medium (η_0) increases. For the same reasons, the positron capture by nuclei may also be ignored.

As regards the possible effect of electron capture by nuclei with decreasing Z and decelerating nucleosynthesis, its magnitude is small for the following reasons. First, the nuclei at such a temperature ($T_9 = 5$) are devoid of electron shells, and, therefore, no κ captute takes place. Second, the electron-capture rate significantly decreases with increasing neutron

excess (see, e.g., Thielemann *et al.* 1986) and, hence, it is small and much less than the rate of β decay for the neutron-rich nuclei under consideration.

Here, it should also be noted that currently available calculations of the rates for weak processes in terms of the shell model for the nucleus (Martinez-Pinedo *et al.* 2000) show that their values are much lower than those calculated previously (Fuller *et al.* 1985), which can change the results of collapse calculations.

RESULTS OF CALCULATIONS

In a series of calculations based on the kinetic model described above, we investigated the nucleosynthesis near the iron peak by taking into account the reactions with neutrons, protons, and β decay at $T_9 = 5$ and $\rho = 10^9$ g cm^{-3} characteristic of type Ia supernovae.

As the initial conditions for nucleosynthesis we chose either the distribution of nuclei in A and Z derived in the approximation of nuclear statistical equilibrium (Panov *et al.* 2001b) or one most abundant nucleus (which had virtually no effect on the final results) with the corresponding neutron excess η_0.

First, we checked the effect of incompleteness of the nuclear data (in our case, the reaction rates) on the yields of elements in nucleosynthesis. Our calculations were performed both using the previously incomplete library of nuclear reaction rates (Fig. 1) and with a fully consistent nuclear data library (Fig. 2). Figure 3 shows the edge effect when the rates of charged-particle reactions for nuclei with $Z > 45$ at $\eta_0 = 0.26$ are disregarded. For $\eta_0 < 0.22$, the results differ only slightly. However, as we see from Figs. 1–3, for an initial neutron excess $\eta_0 \geq 0.24$, new isotopes with $A \geq 130$ are produced through the edge effect caused by incompleteness of the bank of charged-particle reaction rates, an effect which leads to the artificial formation of two regions: in one region, an equilibrium disturbed by a weak leak of matter through β decay into the region with $Z > 45$ is established; and in the other region with $Z > 45$, the charged-particle reactions are frozen and the r-process proceeds. Therefore, for an appreciable neutron excess ($\eta_0 \geq 0.24$), the edge effect shows up clearly, and nucleosynthesis calculations at temperatures $T_9 > 1.5$ must be carried out (as was pointed out by Panov and Nadezhin (1999)) by taking into account the reactions of charged particles with all isotopes involved in nucleosynthesis.

The calculated mass distribution of chemical elements as a function of the neutron excess is shown in Fig. 2. In our calculations, we used nuclear data for isotopes with $Z < 46$ (Thielemann *et al.* 1987) and $Z \geq 46$ (Rauscher and Thielemann 2000). We see from the figure that as the initial neutron excess increases in the equilibrium distribution, the yield of heavy elements increases, the fraction of light elements decreases, and a high fraction of nuclei with atomic masses $\sim$130 is formed at $\eta_0 \geq 0.25$; i.e., a significant number of nuclides with atomic masses $\sim$ 80, 130 can be produced even before the onset of rapid nucleosynthesis. Even for a sufficiently long exposure ($\tau = 2$ s), when η becomes < 0.2, the yield of elements with $A \approx 100$–130 remains high. What actually happens when β decay is taken into account (Fig. 4) if the conditions under consideration (temperature and density) are kept constant for a long time: $\tau_{NSE} \gg \tau_\beta$, where τ_β is the time scale of β decay for the most abundant nuclei? With the passage of time, β decay causes the neutron excess to decrease, and other reactions, such as the (n, α), (n, p), and photonuclear reactions, preventing the production of new nuclei with atomic masses >130, successfully compete with neutron capture. As a result, as the neutrons are depleted and the neutron excess decreases with time, the fractions of nuclei with $A \sim 80, 130$ remain on the same order of magnitude; the yield of nuclei with $80 < A < 130$ decreases by several orders of magnitude; and the yield of nuclei with $A < 80$ increases. Thus, the yields of isotopes under the conditions in question are determined mainly by equilibrium nucleosynthesis. However, for an appreciable nucleosynthesis duration at a high temperature (in our case, at $T_9 \approx 5$) of $\tau \geq 0.5$ s, the abundance peak from the position $A \sim 120$ shifts to $A \sim 130$ because of the change in neutron excess η of the medium through β decay (Fig. 5) and, as a result, in equilibrium.

For $\eta_0 < 0.25$, β decay causes the neutron excess η to decrease for a sufficiently long maintenance of the process ($t \gg \tau_\beta$). Accordingly, the isotopic composition smoothly changes (with time) as the yield of elements with $90 < A < 120$ decrease and the fractions Y_A increase for $A < 70$.

If, however, the temperature decreases by several times and the neutron density is still high enough ($N_n > 10^{24}$), then a successful r-process is possible on the basis of the already formed peaks with $A \sim 80, 130$. It requires a much smaller number of neutrons than the classical r-process, which starts in the range of iron-peak elements. This is clearly seen in Fig. 1 from Lyutostanskii and Panov (1988) and in Figs. 2 and 3 from Lyutostanskii *et al.* (1985). These figures illustrate the velocity of the nucleosynthesis wave, from which it is clear that the formation time of the third peak on the abundance curve is much shorter than the formation time of the second peak.

Note that the degree of initial matter neutronization is of crucial importance in the production of nuclei with $A \sim 130$. Thus, as the neutron excess decreases with expanding and cooling matter, the peak

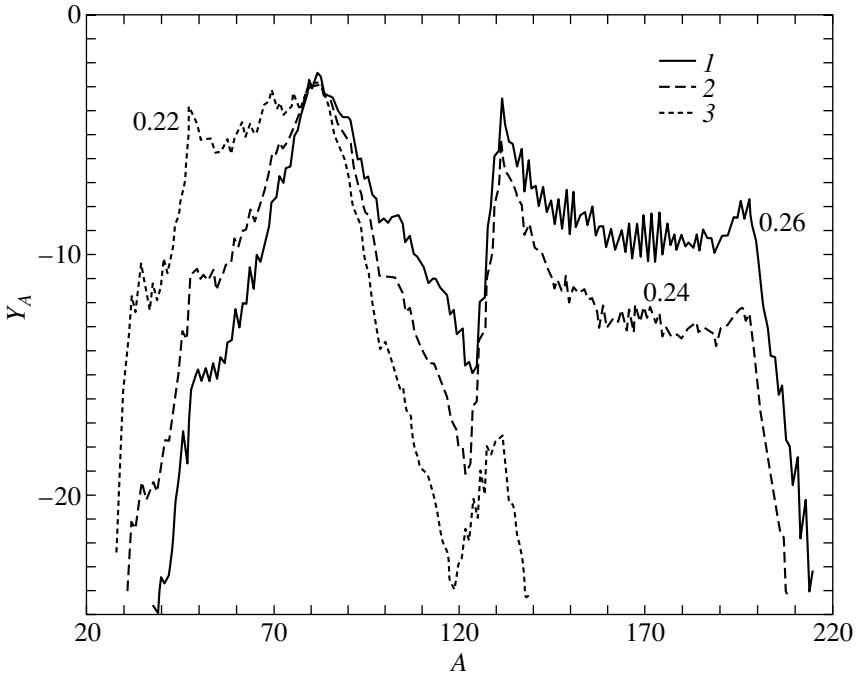

Fig. 1. Calculated mass fractions of the chemical elements produced at $T_9 = 5.0$ and $\rho = 2 \times 10^9$ g cm^{-3} as a function of the initial neutron excess: (*1*) $\eta_0 = 0.26$; (*2*) $\eta_0 = 0.24$; (*3*) $\eta_0 = 0.22$. The initial composition—Fe and neutrons; the charged-particle reactions were taken into account for all nuclei with $Z < 45$ (Thielemann *et al.* 1987). The duration of the calculations is $\tau = 2$ s.

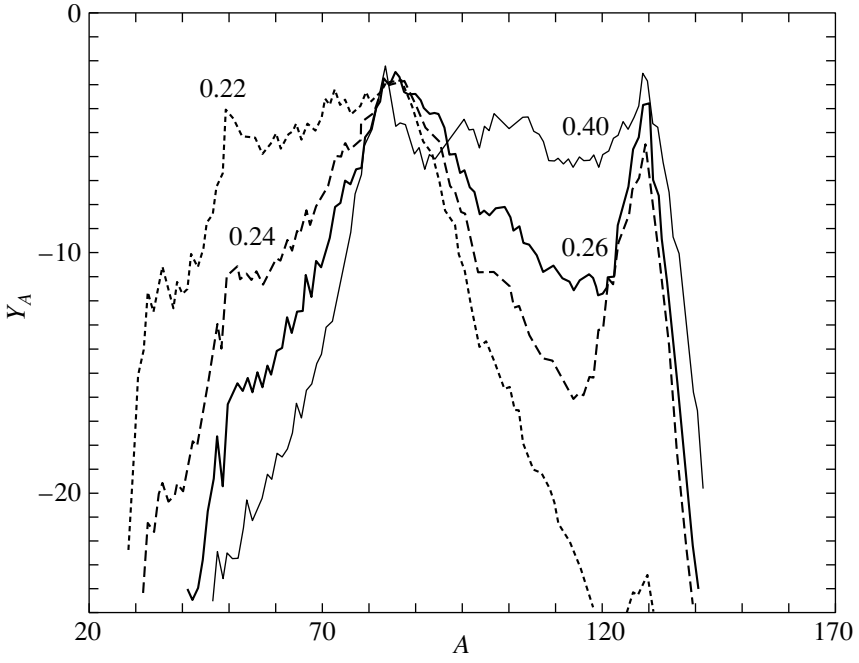

Fig. 2. Same as Fig. 1 but the charged-particle reactions were taken into account up to the nuclei with $Z < 74$ (Rauscher and Thielemann 2000). The initial values of η_0 are indicated near the curves.

at $A \sim 130$ disappears for $\eta_0 < 0.25$ and is preserved for $\eta_0 > 0.25$ (see Fig. 6). As regards the fission processes, they can also contribute in several cases to the formation of the second peak (Panov *et al.* 2001a). Nevertheless, this contribution is apparently minor, except for the cases where the time it takes for the

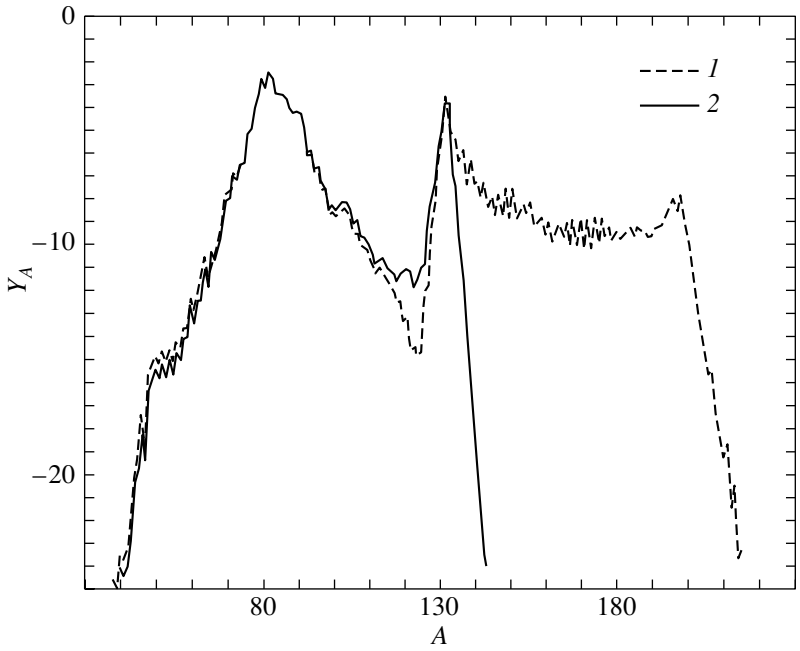

Fig. 3. The calculated yields of elements for $\eta_0 = 0.26$ as a function of the range in which the charged-particle reactions are taken into account. (*1*) The calculated the rates with charged particles based on data from Thielemann *et al.* (1987); (*2*) the rates for elements with $Z > 45$ were taken from Rauscher and Thielemann (2000). The initial conditions are the same as in our previous calculations (Figs. 1 and 2).

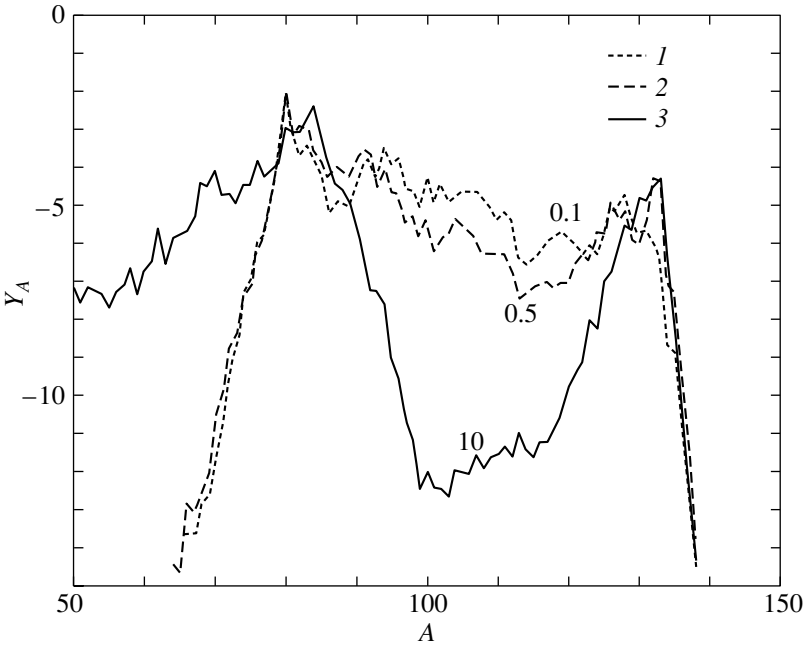

Fig. 4. The mass distribution of isotopes versus nucleosynthesis duration τ for the initial neutron excess $\eta_0 = 0.26$ (the nucleosynthesis time, in seconds, is indicated near the curves). *1*—$\tau = 0.1$ s, *2*—$\tau = 0.5$ s, *3*—$\tau = 10$ s.

nucleosynthesis wave to move from the second peak to the fission region is much shorter than the duration of the r-process (Lyutostanskii *et al.* 1986).

Let us consider an approximate qualitative model for the change in temperature: $T_9 = 5$ at $\tau \leq 0.1$ s; subsequently, the temperature abruptly decreases to $T_9 = 1$ at $\tau > 0.1$ s and remains constant for $\tau = 10$ s. If $\eta_0 = 0.25$, then the number of neutrons is

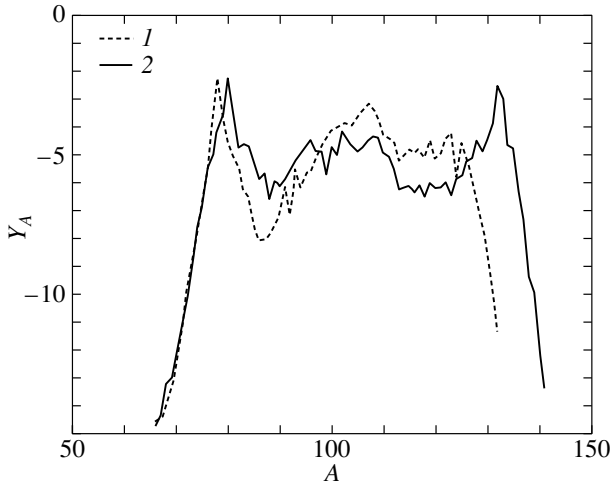

Fig. 5. The yields of elements of different masses as a function of β decay when calculating the nucleosynthesis in a medium with the initial neutron excess $\eta_0 = 0.26$: (*1*) the calculation without β decay, (*2*) with β decay.

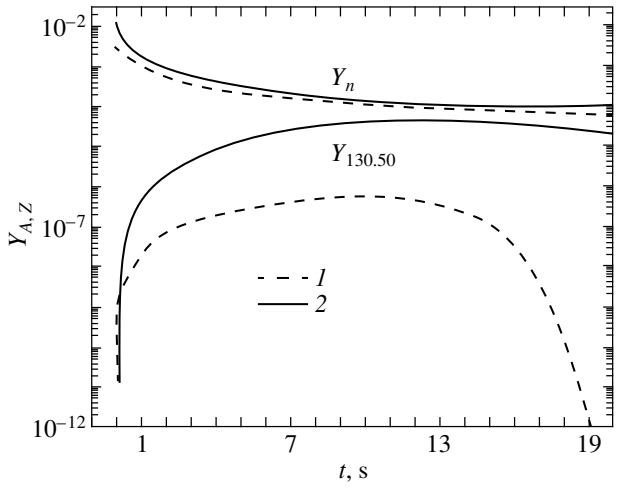

Fig. 6. The production dynamics of nuclides with atomic masses $A \sim 130$ and a change in the number density of free neutrons for various initial neutron excesses: (*1*) $\eta_0 = 0.24$, (*2*) $\eta_0 = 0.26$.

not enough for a successful r-process (Fig. 7). If, however, $\eta_0 = 0.40$, then the number of neutrons is enough for heavier elements to be produced (Fig. 8).

Thus, our results can be formulated as follows:

(1) Our nucleosynthesis calculations with a complete library of nuclear reaction rates (including the reactions not only with neutrons and β decay but also with charged particles) and with a consistent approach to determining all nuclear data show that, at least for the conditions under consideration ($T_9 \approx 5$, $\rho = 10^9$ g cm^{-3}, $\eta_0 > 0.25$), chemical elements are produced up to the peak at $A \sim 130$, and no heavier isotopes are produced, despite allowance for β decay and the high initial density of free neutrons.

(2) As the neutron excess rises, the yield of elements with $A > 80$ in the matter increases significantly. However, the neutron excess decreases with time through β decay, the density of free neutrons drops, and the nucleosynthesis terminates; the produced isotopes with mass A pass from the neutron-rich region to the region of long-lived isotopes with larger Z. If the initial value $\eta_0 < 0.25$, then elements with $A > 100$ burn out with decreasing neutron flux; if, alternatively, the initial value $\eta_0 > 0.25$, then a high yield of elements with $A \sim 130$ is preserved.

(3) When the temperature drops below $T_9 = 5$, the charged-particle reactions decrease sharply in importance. This occurs not only because of the reduction in reaction rates but primarily because of the decrease in the density of protons and α-particles by several orders of magnitude. As a result, the neutrons are rapidly absorbed and the production of new nuclei heavier than $A \sim 80$ decelerates sharply.

(4) As was assumed by Ptitsyn and Chechetkin (1982) and Panov *et al.* (1995), an enhanced neutronization of the expanding matter can give rise to a broad (in mass) spectrum of new nuclei; the isotopes up to $A \approx 120$ are produced through strong and electromagnetic interactions, while the elements of the peak at $A \sim 130$ are formed later, after an equilibrium in charged-particle reactions is established on time scales $t > 0.01$ s through β decay, which causes a decrease in neutron excess and a displacement of the equilibrium to the region of long-lived and stable nuclei. At a large neutron excess ($T_9 = 5$, $0.19 < \eta_0 < 0.3$), the r-process will not proceed even if the temperature decreases, because the fraction of free neutrons is too small.

(5) When an even larger number of free neutrons emerge ($\eta_0 \geq 0.3$), the r-process can be realized even with a large number of seed nuclei, such that the n/Fe ratio is small (< 1). This means that during the explosion and breakup of a low-mass type I supernova, the same conditions as, for example, during neutron-star mergers (Rosswog *et al.* 1999), where part of the neutron-rich, dense and hot matter is ejected into the ambient medium, can arise. If the cooling time from $T_9 \sim 5$ to $T_9 \sim 1.5$ is short ($\sim$0.1 s), then the starting conditions for rapid nucleosynthesis during the cooling to the temperatures typical of the r-process can be close to the ideal ones: there are a large number of seed nuclei with masses from 80 to 130, and the number of free neutrons is still large enough for some of the nuclei to transform into heavier nuclei up to $A \sim 196$ and heavier through multiple neutron captures and β decay.

(6) Based on our results, we refine the definition of the *rbc*-process (Ptitsyn and Chechetkin 1982)

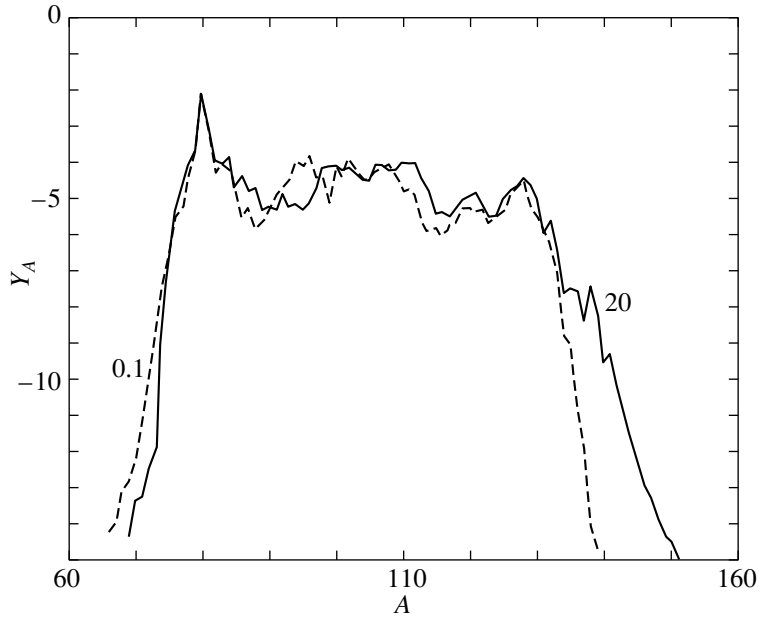

Fig. 7. The *rbc*-process for the initial neutron excess $\eta_0 = 0.26$. For a description of the model for the change of conditions, see the text. The nucleosynthesis duration is indicated near the curves.

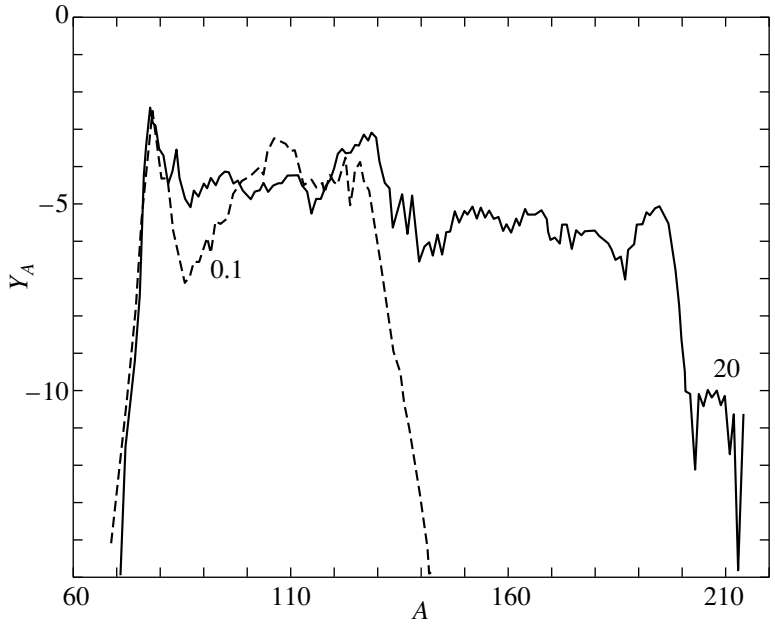

Fig. 8. Same as Fig. 7 for $\eta_0 = 0.40$.

given above. For nuclei with $Z < Z^*$ at a high degree of neutronization ($\eta_0 \leq 0.3$) and at temperatures $T_9 \sim 5$, nuclei with atomic masses $80 < A < 130$ are produced. Even when the neutron density decreases through β decay, if the density and temperature change only slightly during a long period ($\tau \geq 1$ s), the distribution of nuclei Y_A changes little and the elements with $A \sim 80, 130$ remain most abundant; β decay may play a prominent role in forming the peak at $A \sim 130$. As the temperature decreases and as the nuclear reactions with charged particles "are frozen" in the entire range of heavy elements, rapid nucleosynthesis begins (Panov and Nadyozhin 1999). The number of free neutrons is still enough for this nucleosynthesis to be successful. We emphasize once again that, in contrast to the r-

process, far fewer free neutrons are required in the scenario under consideration, because the nuclei with $80 < A < 130$ have already been produced.

Thus, we have shown that the production of nuclei with $A \sim 80$–130 for an enhanced neutron excess in one way or another can solve the question of r-element formation in the following cases: (1) if the initial neutron excess is large and if the hydrodynamic time scale is short, then all elements from iron to uranium can be produced, with the scenario for the production of heavy elements considered here requiring a much smaller number of free neutrons; and (2) if the number of neutrons after matter cooling is not enough for the elements heavier than xenon to be synthesized, then this scenario for the production of chemical elements can be of fundamental importance in elucidating the formation of these elements in nature, because currently available models of the r-process during explosions of high-mass type II supernovae or neutron-star mergers show good agreement with observations only for atomic masses $A > 120$. Therefore, in the scenario for the production of heavy elements under consideration, which assumes a high degree of initial matter neutronization ($\eta_0 \geq 0.30$), an additional r-process responsible for the production of nuclei with $A < 130$ can be realized.

ACKNOWLEDGMENTS

We wish to thank V.S. Imshennik and D.K. Nadyozhin for interest and for valuable remarks and S.I. Blinnikov for helpful discussions of the methods for solving stiff systems of differential equations. This work was supported in part by the Russian Foundation for Basic Research, project nos. 00-02-17230 and 00-02-17253, and the Astronomy Programs.

REFERENCES

1. G. S. Bisnovatyi-Kogan and V. M. Chechetkin, Usp. Fiz. Nauk **127**, 263 (1979) [Sov. Phys. Usp. **22**, 89 (1979)].
2. S. I. Blinnikov and I. V. Panov, Pis'ma Astron. Zh. **22**, 45 (1996) [Astron. Lett. **22**, 39 (1996)].
3. S. I. Blinnikov, I. V. Panov, D. A. Ptitsyn, and V. M. Chechetkin, Pis'ma Astron. Zh. **21**, 872 (1995) [Astron. Lett. **21**, 786 (1995)].
4. G. R. Burbidge, E. M. Burbidge, W. A. Fowler, and F. Hoyle, Rev. Mod. Phys. **29**, 547 (1957).
5. A. G. W. Cameron, Astrophys. Space Sci. **82**, 123 (1982).
6. V. M. Chechetkin, S. S. Gerstein, V. S. Imshennik, *et al.*, Astrophys. Space Sci. **67**, 61 (1980).
7. J. J. Cowan, F.-K. Thielemann, and J. W. Truran, Phys. Rep. **208**, 267 (1991).
8. R. I. Epstein, S. A. Colgate, and W. C. Haxton, Phys. Rev. Lett. **61**, 2038 (1988).
9. W. Fowler and F. Hoyle, *Nucleosynthesis in Massive Stars and Supernovae* (Univ. of Chicago Press, Chicago, 1965; Mir, Moscow, 1967).
10. C. Freiburghaus, S. Rosswog, and F.-K. Thielemann, Astrophys. J. Lett. **525**, L121 (1999a).
11. C. Freiburghaus, J.-F. Rembges, T. Rausher, *et al.*, Astrophys. J. **516**, 381 (1999b).
12. G. M. Fuller, W. A. Fowler, and M. J. Newman, Astrophys. J. **293**, 1 (1985).
13. E. R. Hilf, H. V. Groote, and K. Takahashi, CERN Report No. CERN-76-13 (1976), p. 142.
14. W. Hillebrandt, Space Sci. Rev. **21**, 639 (1978).
15. W. Hillebrandt and F.-K. Thielemann, Astron. Astrophys. **58**, 357 (1977).
16. W. Hillebrandt, K. Nomoto, and R. G. Wolf, Astron. Astrophys. **133**, 175 (1984).
17. V. S. Imshennik, Pis'ma Astron. Zh. **18**, 489 (1992) [Sov. Astron. Lett. **18**, 194 (1992)].
18. V. S. Imshennik, N. L. Kal'yanova, A. V. Koldoba, and V. M. Chechetkin, Pis'ma Astron. Zh. **25**, 250 (1999) [Astron. Lett. **25**, 206 (1999)].
19. L. N. Ivanova, V. S. Imshennik, and V. M. Chechetkin, Astron. Zh. **54**, 354 (1977) [Sov. Astron. **21**, 197 (1977)].
20. L. N. Ivanova, V. S. Imshennik, and V. M. Chechetkin, Preprint No. 109 (Inst. Teor. Éksp. Fiz., 1983).
21. F. Käppeler, H. Beer, and K. Wisshak, Rep. Prog. Phys. **52**, 945 (1989).
22. F. Käppeler, F.-K. Thielemann, and M. Wiesher, Annu. Rev. Nucl. Part. Sci. **48**, 175 (1998).
23. K.-L. Kratz, J.-P. Bitouzet, F.-K. Thielemann, *et al.*, Astrophys. J. **403**, 216 (1993).
24. K. R. Lang, *Astrophysical Formulae: A Compendium for the Physicist and Astrophysicist* (Springer-Verlag, Berlin, 1974; Mir, Moscow, 1978).
25. J. M. Lattimer and D. N. Schramm, Astrophys. J. Lett. **192**, L145 (1974).
26. J. M. Lattimer and D. N. Schramm, Astrophys. J. **210**, 549 (1976).
27. Yu. S. Lyutostanskiĭ and I. V. Panov, Pis'ma Astron. Zh. **14**, 70 (1988) [Sov. Astron. Lett. **14**, 70 (1988)].
28. Yu. S. Lyutostanskiĭ, D. A. Ptitsyn, O. N. Sinyukova, *et al.*, Yad. Fiz. **42**, 215 (1985) [Sov. J. Nucl. Phys. **42**, 136 (1985)].
29. Yu. S. Lyutostanskiĭ, I. V. Panov, O. N. Sinyukova, *et al.*, Yad. Fiz. **44**, 66 (1986) [Sov. J. Nucl. Phys. **44**, 43 (1986)].
30. G. Martinez-Pinedo, K. Langanke, and D. J. Dean, Astrophys. J., Suppl. Ser. **126**, 493 (2000).
31. G. J. Mathews and J. J. Cowan, Nature **345**, 491 (1990).
32. P. Möller and J. R. Nix, Nucl. Phys. A **536**, 20 (1992).
33. P. Möller, J. R. Nix, and K.-L. Kratz, At. Data Nucl. Data Tables **66**, 131 (1997).
34. D. K. Nadyozhin and I. V. Panov, Pis'ma Astron. Zh. **27**, 516 (2001) [Astron. Lett. **27**, 440 (2001)].
35. D. K. Nadyozhin, I. V. Panov, and S. I. Blinnikov, Astron. Astrophys. **335**, 207 (1998).
36. I. V. Panov and D. K. Nadyozhin, Pis'ma Astron. Zh. **25**, 435 (1999) [Astron. Lett. **25**, 369 (1999)].

37. I. V. Panov, D. A. Ptitsyn, and V. M. Chechetkin, Pis'ma Astron. Zh. **21**, 206 (1995) [Astron. Lett. **21**, 185 (1995)].
38. I. V. Panov, S. Freiburghaus, and F.-K. Thielemann, Nucl. Phys. A **688**, 587c (2001a).
39. I. V. Panov, S. I. Blinnikov, and F.-K. Thielemann, Pis'ma Astron. Zh. **27**, 279 (2001b) [Astron. Lett. **27**, 239 (2001b)].
40. D. A. Ptitsyn and V. M. Chechetkin, Pis'ma Astron. Zh. **8**, 600 (1982) [Sov. Astron. Lett. **8**, 322 (1982)].
41. Y.-Z. Qian, P. Vogel, and G. J. Wasserburg, Astrophys. J. **494**, 285 (1998).
42. T. Rauscher and F.-K. Thielemann, At. Data Nucl. Data Tables **75**, 1 (2000).
43. S. Rosswog, M. Liebendörfer, F.-K. Thielemann, *et al.*, Astron. Astrophys. **341**, 499 (1999).
44. C. Sneden, J. J. Cowan, I. I. Ivans, *et al.*, Astrophys. J. Lett. **533**, L139 (2000); astro-ph/0003086.
45. E. M. D. Symbalisty and D. N. Schramm, Astrophys. Lett. **22**, 143 (1982).
46. K. Takahashi, J. Witti, and H.-Th. Janka, Astron. Astrophys. **286**, 857 (1994).
47. F.-K. Thielemann, K. Nomoto, and K. Yokoi, Astron. Astrophys. **158**, 17 (1986).
48. F.-K. Thielemann, M. Arnould, and W. Truran, in *Advances in Nuclear Astrophysics,* Ed. by E. Vangioni-Flam *et al.* (Editions Frontieres, Gif-sur-Yvette, 1987), p. 525.
49. J. W. Truran, J. J. Cowan, and A. G. W. Cameron, Astrophys. J. Lett. **222**, L63 (1978).
50. G. J. Wasserburg, M. Busso, and R. Gallino, Astrophys. J. Lett. **466**, L109 (1996).
51. J. G. Wheeler, J. J. Cowan, and W. Hillebrandt, Astrophys. J. Lett. **493**, L101 (1998).
52. J. Witti, H.-Th. Janka, K. Takahashi, and W. Hillebrandt, in *Nuclei in the Cosmos,* Ed. by F. Käppeler and K. Wisshak (Inst. of Physics Publ., Bristol, 1993), p. 601.
53. S. E. Woosley and R. D. Hoffman, Astrophys. J. **395**, 202 (1992).
54. S. E. Woosley, J. R. Wilson, G. J. Mathews, *et al.*, Astrophys. J. **433**, 229 (1994).

Translated by V. Astakhov

Astronomy Letters, Vol. 28, No. 7, 2002, pp. 488–493. Translated from Pis'ma v Astronomicheskiĭ Zhurnal, Vol. 28, No. 7, 2002, pp. 554–560.

On the Acceleration of Solar-Flare Charged Particles in a Collapsing Magnetic Trap with an Electric Potential

V. A. Kovalev[1]* and B. V. Somov[2]

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

[2]*Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119899 Russia*

Received January 17, 2002

Abstract—We consider the kinetic problem of charged-particle acceleration in a magnetic trap with converging magnetic mirrors. We show that for a positive electrostatic potential of the trap plasma relative to the mirrors, the efficiency of confinement and acceleration increases for electrons and decreases for ions.

Key words: *solar flares, particle acceleration, electric potential, loss region*

INTRODUCTION

Solving the problem of charged-particle acceleration is of importance in understanding many questions related to the rapid transformation of magnetic energy into the thermal and kinetic energy of particles in cosmic plasma (Somov 2000). The response of cosmic plasma to the effect of accelerated particles is studied by their manifestation in various ranges of electromagnetic radiation, from gamma rays to radio waves. In particular, the observed fine structure of the radio emission from solar flares suggests a complex pattern of the nonequilibrium kinetic processes that take place in plasma. To faithfully describe them, it is important to have a clear idea of the formation and behavior of the accelerated-particle distribution function.

One of the sources of accelerated particles in solar flares can be collapsing magnetic traps (Somov 1994). When the front of a traveling shock wave crosses a magnetic flux tube, magnetic mirrors with a field discontinuity emerge, a magnetic trap with a decreasing length is formed, and its evolution has a collapsing pattern; i.e., its length tends to zero (Somov and Kosugi 1997). In this case, the particle energy increases through the Fermi acceleration mechanism (Fermi 1954).

The solutions for the distribution function of nonrelativistic accelerated electrons in a magnetic trap with magnetic mirrors moving toward each other in the absence of an electric field and Coulomb collisions were obtained by Gisler and Lemons (1990). As applied to solar conditions, the acceleration of relativistic particles in a collapsing trap in the absence of an electric field was considered by Bogachev and Somov (1999). Meanwhile, the presence of a field significantly affects the conditions of particle escape from the trap and, naturally, the acceleration pattern.

In general, investigating the problem requires solving a self-consistent system of kinetic equations for various plasma components in magnetic and electric fields with the inclusion of Coulomb collisions. In this paper, we consider only some aspects of the kinetic problem related to the effect of an electric potential on the particle acceleration and escape conditions.

CHARGED PARTICLES IN A MAGNETIC TRAP

We assume that the magnetic loop rapidly moving from the magnetic reconnection region in a solar flare may be considered as an open trap in which the magnetic field is uniform in most of it but increases sharply from B_1 to B_2 in the mirrors (Fig. 1). The quantity B_2/B_1 is called the mirror ratio. Open traps are known (Pastukhov 1984) to have a high admissible plasma β (the ratio of plasma gas pressure to magnetic pressure) and stability against rough hydrodynamic disturbances.

The length of a dynamical trap with mirrors moving at velocity v_m varies as

$$l = L/L_0 = 1 - t/t_0, \qquad (1)$$

where L_0 is the initial length of the trap at its generation time $t = 0$ and $t_0 = L_0/2v_m$ is the collapse time. Below, instead of t, we use the variable l that

*E-mail: vkovalev@izmiran.troitsk.ru

1063-7737/02/2807-0488$22.00

characterizes the trap contraction. In the absence of Coulomb collisions, there is a transverse adiabatic invariant for the fast particles inside the trap: $v_\perp^2/B = \text{const}$. When moving into a region with a stronger field (into a mirror), the velocity component $v_\perp$ transverse to the magnetic field increases. At the same time, the transverse velocity is conserved between reflections from the mirrors:

$$v_\perp = v_{\perp 0} = \text{const}.$$

The last two conditions do not contradict each other, because the field discontinuity takes place only near the mirrors, where the particle spends a short time compared to the time of its flight in the main homogeneous part of the tube. In other words, we consider the problem in the approximation of a long trap with short mirrors (Somov 1994). This condition is violated near the time of collapse, when the trap length tends to zero.

At the same temperature, the thermal velocities ($V_T = (2kT/m)^{1/2}$) of electrons and protons differ by a factor of $(m_p/m_e)^{1/2}$. Thus, for the hottest flare plasma with $T_e \simeq T_i \simeq 10^8$ K, the thermal electron velocity is $V_{T_e} \sim 5 \times 10^9$ cm s^{-1} and the proton velocity is $V_{T_i} \sim 1.2 \times 10^8$ cm s^{-1}. The mirror velocity is $v_m < 10^8$ cm s^{-1}. For electrons, the condition

$$v_m \ll V_{T_e} \tag{2}$$

is satisfied and the acceleration may be considered as a continuous adiabatic process, which can be written in differential form. After each reflection from a magnetic mirror, the longitudinal velocity component increases by $\delta v_{||} = 2v_m$. The time between two sequential reflections is δt is $L/(v_{||} + v_m)$. Hence,

$$\frac{dv_{||}}{dt} \simeq \frac{\delta v_{||}}{\delta t} = 2v_m \frac{(v_{||} + v_m)}{L} \simeq 2v_m \frac{v_{||}}{L}.$$

Using the variable l, we write the condition for the longitudinal adiabatic invariant:

$$v_{||} l = \text{const}. \tag{3}$$

If, however, the adiabaticity condition (2) is not satisfied, then we must take into account the discrete pattern of acceleration in the trap, and the longitudinal component of the particle momentum increases during an elastic collision with mirrors moving at velocity v_m as

$$v_{||} = v_{||0} + 2v_m \mathcal{N}(x). \tag{4}$$

Here, $\mathcal{N}$ is the number of reflections from the mirrors, which can be determined by using the last two relations:

$$\mathcal{N}(x) = \mathcal{N}\left(\frac{v_{||0}(1-l) + v_m l}{2v_m l}\right), \tag{5}$$

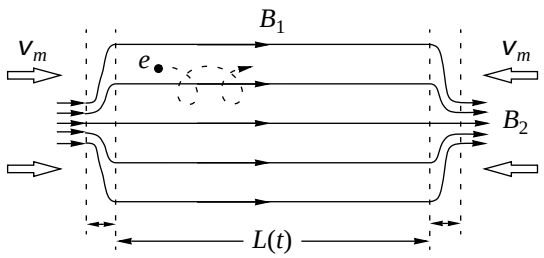

Fig. 1. An idealized model of a long trap with mirrors moving toward each other.

where $\mathcal{N}(x)$ is the integer part of the argument x. The velocity increment does not depend on the particle mass and charge but depends only on the mirror velocity v_m, which may be assumed to be constant. In the absence of an electric field and Coulomb collisions, the trapped particles oscillate with an increasing frequency $(v_{||}/L)$ while repeatedly reflecting from the converging mirrors and significantly increase their kinetic energy.

Since only $v_{||}$ increases, the particle pitch angle decreases, and after a time, such a particle with

$$v_{||} + v_m \geq v_{\perp 0} R, \quad R = (B_2/B_1 - 1)^{1/2} \tag{6}$$

becomes an escaping one, falling into the loss cone that satisfies the condition

$$\sin^2\theta \leq \sin^2\theta_{\text{cr}} = B_1/B_2. \tag{7}$$

The time of particle escape from the trap can be obtained from (6):

$$l_{\text{ex}} = \frac{v_{||0}}{v_{\perp 0} R}. \tag{8}$$

The latter expression is meaningful when its right-hand side takes on values in the interval [0, 1]; otherwise, the particles are not trapped from the very outset ($l = 1$).

The criterion for the necessity of allowing for Coulomb collisions is determined by the ratio of the time of flight of a probe particle (electron) in the trap $\tau = L/v_{||}$ to the time of Coulomb collisions with particles (electrons and ions) of the background plasma τ_s

$$\tau_s = \sqrt{m_e} T^{3/2}/\sqrt{2}\pi n e^4 \Lambda.$$

The collisions may be ignored if $\tau \ll \tau_s$ or $v_{0||} \gg L_0 l^2/\tau_s$. Let us obtain estimates for a flare loop with the particle number density $n = 2 \times 10^9$ cm^{-3} and temperature $T = 10^8$ K. The Coulomb logarithm Λ can vary over the range 10–20. Then, $\tau_s \simeq$ (3–6) s. At the same time, the time of flight of thermal electrons in a trap with $L_0 = 2 \times 10^9$ cm is considerably shorter ($\simeq$0.4 s); therefore, the collisionless condition is satisfied.

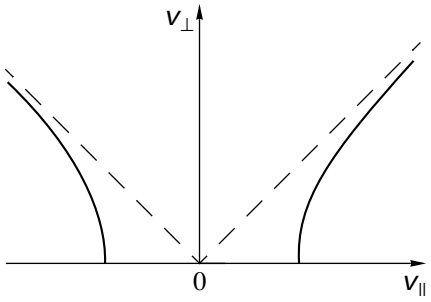

Fig. 2. The boundary of the electron-confinement region in velocity space. The dashed line indicates the loss cone (Pastukhov 1984).

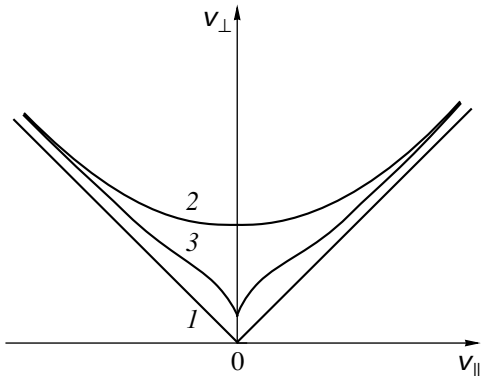

Fig. 3. The boundary of the ion-confinement region; (*1*) cone, (*2*) hyperboloid, (*3*) a self-consistent boundary (Pastukhov 1984).

ON THE ROLE OF AN ELECTROSTATIC POTENTIAL

The presence of an electrostatic potential significantly changes the conditions of particle escape from the trap. The electric fields produced in the traps under consideration can have distinctly different origins. For instance, they can result from the separation of volume electric charge due to the different behavior of electrons and ions inside the trap. Another cause can generally be any deviation of the magnetic field from a potential (i.e., current-free) field. Any currents, even in a force-free magnetic configuration, require electric fields for their maintenance. For dynamical magnetic traps, the currents associated with the forces acting on plasma must be present.

At the kinetic level, the origins of the electric field can be different. For example, the classical losses are determined by Coulomb scattering into the loss cone. In this case, ions are scattered approximately a factor of $(m_i/m_e)^{1/2}$ (where m_i and m_e are the ion and electron masses, respectively) more slowly than electrons with the same energy. To equalize the losses of ions and electrons and, thereby, to conserve its quasineutrality, the plasma acquires a positive potential relative to the mirrors (Pastukhov 1984). The electric potential can also result from a difference in anisotropy of the electron and ion distributions (Spicer and Emslie 1990). In general, the electric potential is the solution of a self-consistent problem.

Without detailing the nature of the electric fields in the trap, we consider their effect on the dynamics of the accelerated particles in general form by characterizing this effect by some potential φ, positive or negative. For simplicity, we assume here that the potential, as well as the magnetic field, changes abruptly at the trap ends but does not depend on l; there is no electric field in the main part of the trap.

For a positive plasma potential φ relative to the mirrors, the energy of electric charge $Ze\,(e>0)$ is

$$\mathcal{E} = mv^2/2 + Ze\varphi. \qquad (9)$$

The region of trapped electrons ($Z=-1, \mathcal{E}_\parallel < 0$, a potential well) in velocity space enlarges compared to the ordinary cone and takes the shape of a two-sheet hyperboloid (Fig. 2):

$$v_{||0}^2 < v_{\perp 0}^2 R^2 + v_{\varphi e}^2, \quad v_{\varphi e}^2 = 2e\varphi/m_e. \qquad (10)$$

The minimum velocity of electron escape from the trap is

$$v_{\min} = (2e\varphi/m_e)^{1/2};$$

thus, the trapping and acceleration efficiencies increase.

A positive electric potential also results in the confinement of electrons in the trap; the time at which they fall into the loss hyperboloid is

$$l_{\rm ex}^e = \frac{v_{||0}}{\sqrt{v_{\perp 0}^2 R^2 + v_{\varphi e}^2}}. \qquad (11)$$

The case where the right-hand side of the latter expression is >1 corresponds to the nontrapped particles at time $l=1$.

The situation for ions ($Z=1$, $\mathcal{E}_\parallel > 0$, a potential barrier) is different. The region of trapped particles shrinks compared to the cone and takes the shape of a one-sheet hyperboloid (Fig. 3):

$$R^2 v_{\perp 0}^2 > v_{||0}^2 + v_{\varphi i}^2, \quad v_{\varphi i}^2 = 2Ze\varphi/m_i. \qquad (12)$$

In this case, the trap region in velocity space decreases with increasing Z and grows with m_i. Thus, a positive potential reduces the efficiency of ion confinement and acceleration. The time of their escape into the loss hyperboloid is

$$l_{\rm ex}^i = \frac{v_{||0}}{\sqrt{v_{\perp 0}^2 R^2 - v_{\varphi i}^2}}. \qquad (13)$$

The case where the radicand is <0 or $l_{\rm ex}^i > 1$ corresponds to the particles that are already in the loss

region at the initial time. The condition for the critical angle (7) in the presence of an electric field is modified (Spicer and Emslie 1990):

$$\sin^2\theta_{\rm cr} = \left(1 \pm \frac{v_\varphi^2}{v^2}\right) B_1/B_2. \qquad (14)$$

The plus and minus signs refer to ions and electrons, respectively. The right-hand side of Eq. (14) for ions can be >1; in this case, there can be no ions with such velocities. On the other hand, for electrons, the right-hand side is <0, suggesting that the particles are trapped. Thus, a positive potential confines the relatively slow electrons in the trap and expels ions. A dependence of the critical angle on the particle velocity appears: the critical angle decreases for ions and increases for electrons.

CALCULATION OF THE DISTRIBUTIONS

Because of the presence of mirrors, the particles are separated into trapped and escaping ones: the distribution segments corresponding to the particles that satisfy the escape conditions are cut out from the distribution of the particles injected into the trap. These particles precipitate into the loss region and are not involved in the acceleration. Therefore, the initial distribution $f(1)$ corresponding to $l = 1$ generally differs from the injected distribution f_0.

The total number of trapped particles is (Bogachev and Somov 1999)

$$N(l) = 4\pi N_0 \int\limits_{\mathcal{D}}\int f_0(v_{0||}, v_{0\perp}) v_{0\perp} dv_{0||} dv_{0\perp}, \quad (15)$$

where $\mathcal{D}$ is the domain of integration. We took into account the fact that the longitudinal-velocity distribution is symmetric in a symmetric trap. The change in particle density $n(l)$ is determined by the change in the total number of particles $N(l)$ and by the change in the trap volume. For a trap with a constant cross section, the volume is proportional to $L = L_0/l$. Then,

$$n(l) = \frac{N(l)}{L} = \frac{n_0}{N_0}\frac{N(l)}{l}. \qquad (16)$$

The Solution for Trapped Charged Particles for $\varphi = 0$

In this simplest case, the cone (6) corresponds to the trapped particles (both electrons and ions, irrespective of the charge):

$$0 \le v_{0||} < Rlv_{0\perp}, \quad 0 \le v_{0\perp} < \infty. \qquad (17)$$

For the distribution function, we have (Bogachev and Somov 1999):

$$f(v_{||}, v_\perp, l) = \frac{N_0 l}{N(l)}\Theta(Rv_\perp - v_{||}) f_0(lv_{||}, v_\perp), \quad (18)$$

where $\Theta(x \ge 0) = 1,\ \Theta(x < 0) = 0$.

As the function of injected particles f_0 we use the Maxwellian distribution

$$f_0(v_{||}\, v_\perp) = \frac{1}{\pi^{3/2} V_T^3} \exp(-v^2/V_{\rm T}^2). \qquad (19)$$

The sought distribution (18) becomes a two-temperature one, $(V_{T\perp} = V_T, V_{T||} = V_T/l)$.

Then, integrating (19) over the domain (17) yields

$$N(l)/N_0 = ql, \qquad (20)$$

$$n(l)/n_o = q, \qquad (21)$$

$$q = \frac{R}{\sqrt{1 + R^2 l^2}}.$$

For $l = 1$, the coefficient q characterizes the extent to which the injected flare particles are trapped. An increase in R causes the trap region to increase (the loss cone to shrink). Therefore, the flux of the particles precipitating into the loss cone determined by the derivative $dN(l)/dl$ decreases with increasing R at an early stage of trap contraction but increases at a late stage.

The Solution for Captured Electrons for a Positive Potential $\varphi > 0$

According to (10), the domain of integration is

$$0 \le v_{0||} < v_{*e}, \quad 0 \le v_{0\perp} < \infty,$$
$$v_{*e} = l(v_{o\perp}^2 R^2 + v_{\varphi e}^2)^{1/2}.$$

Then, instead of (18), we have

$$f_e(v_{||}, v_\perp, l) = \frac{N_0 l}{N(l)}\Theta(Rv_\perp - v_{*e}) f_0(lv_{||}, v_\perp), \quad (22)$$

$$N_e(l) = 4\pi N_0 \int\limits_0^\infty v_{0\perp} dv_{0\perp} \int\limits_0^{v_*} f_0(v_{0||}, v_{0\perp}) dv_{0||}. \qquad (23)$$

The latter relation can be written as

$$\frac{N_e(l)}{N_0} = \frac{4}{\sqrt{\pi}} \int\limits_0^\infty y \exp(-y^2) dy J_1(y), \qquad (24)$$

$$J_1(y) = \int\limits_0^{u_*(y)} \exp(-x^2) dx,$$

where

$$x = v_{0||}/V_{T_e}, \quad y = v_{0\perp}/V_{T_e},$$
$$u_{\varphi e} = v_{\varphi e}/V_{T_e}, \quad u_{*e} = v_{*e}/V_{T_e}.$$

We integrate expression (24) by parts:

$$J_1(y) = t, \quad dt = \frac{du_*(y)}{dy}\exp(-u_*^2(y))dy,$$

$$dz = y\exp(-y^2)dy, \quad z = -\frac{1}{2}\exp(-y^2).$$

As a result, given

$$J_1(0) = \frac{\sqrt{\pi}}{2}\mathrm{erf}(lu_{\varphi e})$$

and

$$\frac{1}{\sqrt{\pi}}\int\limits_0^\infty \frac{\exp(-\xi)}{(\xi+a)^{1/2}}d\xi = e^a(1-\sqrt{a}),$$

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int\limits_0^x \exp(-\xi^2)d\xi,$$

$$a = \frac{u_{\varphi e}^2}{q^2}, \quad \xi = y\sqrt{1+R^2l^2},$$

we obtain the sought relations:

$$\frac{N_e(l)}{N_o} = ql\exp\left(\frac{u_{\varphi e}^2}{R^2}\right) \tag{25}$$

$$\times\left[1-\mathrm{erf}\left(\frac{u_{\varphi e}}{q}\right)\right] + \mathrm{erf}(lu_{\varphi e}).$$

For a weak potential $u_{\varphi e} \ll 1$, using

$$\mathrm{erf}(x) \simeq \frac{2}{\sqrt{\pi}}x,$$

we have

$$\frac{N_e(l)}{N_0} \simeq ql + \frac{2}{\sqrt{\pi}}lu_{\varphi_e}, \tag{26}$$

$$\frac{n_e(l)}{n_0} \simeq q + \frac{2}{\sqrt{\pi}}u_{\varphi_e}. \tag{27}$$

We see that the presence of a positive potential causes the electron trapping efficiency to increase.

The Solution for Captured Ions for a Positive Potential $\varphi > 0$

According to (13), the domain of integration is

$$0 \le v_{0||} < v_{*i}, \quad v_{\varphi i} < v_{0\perp} < \infty,$$

where

$$v_{*i} = l(v_{0\perp}^2R^2 - v_{\varphi i}^2)^{1/2}.$$

Then, instead of (18), we have

$$f_i(v_{||}, v_\perp, l) = \frac{N_0 l}{N(l)}\Theta(Rv_\perp - v_{||*})f_0(lv_{||}, v_\perp) \tag{28}$$

$$N_i(l) = 4\pi N_0\int\limits_{v_{\varphi i}}^\infty v_{0\perp}dv_{0\perp}\int\limits_0^{v_{*i}} f_0(v_{0||}, v_{0\perp})dv_{0||}. \tag{29}$$

Here,

$$u_{\varphi i}^2 = \frac{v_{\varphi i}^2}{V_{T_i}^2} = \frac{e\varphi}{kT_i}.$$

Repeating the same operations as for electrons, we obtain

$$\frac{N_i(l)}{N_0} = \exp(-u_{\varphi i}^2)\mathrm{erf}(lru_{\varphi i}) \tag{30}$$

$$+ ql\exp(l^2u_{\varphi i}^2)\frac{2}{\sqrt{\pi}}\int\limits_b^\infty \xi\frac{\exp(-\xi^2)}{(\xi^2-a^2)^{1/2}}d\xi,$$

where

$$r = \sqrt{R^2-1},\ a = \frac{u_{\varphi i}}{q},\ b = u_{\varphi i}\sqrt{1+R^2l^2},$$

$$\xi = \frac{v_{o\perp}}{V_T}\sqrt{1+R^2l^2}.$$

In contrast to the previous case, the integral in (30)

$$J(u) = \int\limits_{b(u)}^\infty f(\xi, u)d\xi,$$

where

$$f(\xi, u) = \xi\frac{\exp(-\xi^2)}{(\xi^2-a^2)^{1/2}},$$

cannot be calculated analytically; however, in the case of interest, $u_{\varphi i} \ll 1$, using the expansion at $u = 0$

$$J(u) \simeq J(0) + \left.\frac{dJ(u)}{du}\right|_{u=0} u,$$

where the derivative is

$$\frac{dJ(u)}{du} = \int\limits_{b(u)}^\infty \frac{df(\xi, u)}{du}d\xi - \frac{db(u)}{du}f(\xi, u)|_{\xi=b(u)},$$

we obtain

$$\frac{N_i(l)}{N_0} \simeq ql - \frac{2lR^2}{\sqrt{\pi}r}u_{\varphi i},$$

$$\frac{n_i(l)}{n_0} \simeq q - \frac{2R^2}{\sqrt{\pi}r}u_{\varphi i}.$$

CONCLUSIONS

Thus, we have solved the kinetic problem of charged-particle acceleration in a collapsing magnetic trap with a given electric potential and converging mirrors. We derived analytic formulas for the

total number of particles and their number density as a function of the trap contraction parameter. We showed that for a positive trap plasma potential relative to the mirrors, the efficiency of confinement and acceleration of electrons and their flux into the loss region increase, whereas those for positive ions decrease compared to the case of $\varphi = 0$, in which the Fermi acceleration efficiency does not depend on the charge.

Here, the competing processes must be taken into account. As was pointed out above, on the one hand, electrons, being more mobile than ions, are scattered faster into the loss region because of Coulomb collisions. On the other hand, the region of ion escape in velocity space grows and the efficiency of their acceleration decreases. In this case, the plasma can generally acquire a negative potential, under which the ion confinement and acceleration is more efficient. Plasma quasineutrality can be achieved when the electron and ion fluxes are equal.

In general, determining the exact self-consistent boundary of particle confinement is a more complex problem than the limiting cases of the cone and hyperboloids considered here (Pastukhov 1984). We restricted our analysis to an isotropic distribution of the particles injected into the trap. Meanwhile, anisotropic nonuniform distributions are most probable. In addition, for high-temperature turbulent current sheets, the electrons and ions can have different temperatures with $T_e \gg T_i$ (Somov 2000).

Bases on analysis of the problem, we can establish that even in the collisionless case with isotropic injection, an electric potential can emerge, because the electron mobility is higher than the ion mobility. Even at the time of trap formation, the particles are separated into trapped and flyby ones. The flyby electrons leave the trap faster; the trap plasma acquires a positive potential, expelling protons. A plasma distribution with an electric potential is established when plasma quasineutrality is provided. Plasma anisotropy leads to the growth of electromagnetic instabilities (Mikhaĭlovskiĭ 1975).

The model of charged-particle acceleration in collapsing magnetic traps under discussion can be considered in two aspects. On the one hand, such traps emerge when a magnetic loop interacts with a shock wave (Gisler and Lemons 1990). On the other hand, collapsing traps are an inevitable consequence of reconnection in the corona. To be more precise, they represent the part of the reconnected magnetic fluxes that move downward, toward the chromosphere and photosphere. From the viewpoint of particle acceleration theory, traps are an important detailing of the current-sheet model by S.I. Syrovatskiĭ. The high-energy particles (SCR) accelerated by the electric field in a current sheet escape into interplanetary space along open field lines. At the same time, the thermal particles heated in a current sheet and low-energy particles can be trapped and additionally accelerated in them to significant energies, as our calculations show.

These problems will be further discussed in our forthcoming papers.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 99-02-16076 (V.A.K.) and 99-02-16344 (B.V.S.).

REFERENCES

1. S. A. Bogachev and B. V. Somov, Izv. Akad. Nauk, Ser. Fiz. **63**, 1555 (1999).
2. E. Fermi, Astrophys. J. **119**, 1 (1954).
3. G. Gisler and D. Lemons, J. Geophys. Res. **95** (A9), 14925 (1990).
4. A. B. Mikhailovskiĭ, *Theory of Plasma Instabilities* (Atomizdat, Moscow, 1975; Consultants Bureau, New York, 1974), Vol. 1.
5. V. P. Pastukhov, in *Reviews of Plasma Physics,* Ed. by B. B. Kadomtsev (Énergoatomizdat, Moscow, 1984; Consultants Bureau, New York, 1987), Vol. 13.
6. B. V. Somov, in *Fundamentals of Cosmic Electrodynamics* (Kluwer, Dordrecht, 1994), p. 364.
7. B. V. Somov, in *Cosmic Plasma Physics* (Kluwer, Dordrecht, 2000), p. 652.
8. B. V. Somov and T. Kosugi, Astrophys. J. **485**, 859 (1997).
9. D. S. Spicer and A. G. Emslie, Astrophys. J. **330**, 997 (1988).

Translated by G. Rudnitskiĭ