Baryon asymmetry of the universe

A. D. Sakharov

P.N. Lebedev Physics Institute, Academy of Sciences of the USSR
Usp. Fiz. Nauk **161**, 110–120 (May 1991)

(Review paper presented at the A. A. Friedmann Centenary Conference, Leningrad, 22–26 June, 1988)[1)]

The concept of a nonstationary Universe whose foundation was laid by Friedmann is of a vast scientific and philosophical significance.

Among the problems, whose very formulation was impossible in the pre-Friedmann time there is the problem of baryon asymmetry of the Universe. How can one explain why there are only baryons in the observable part of the Universe, and not antibaryons? What determines quantitatively the magnitude of the baryonic asymmetry (the latter is conventionally characterized by the ratio

$$(\mathrm{BAU}) = \frac{n_\mathrm{B}}{n_\gamma} \sim 10^{-9};$$

where n_B is the mean density of baryons in the Universe, n_γ is the relic photon density, in the order of magnitude equal to the density of entropy). The value of BAU is known but approximately because of: (1) the inaccuracy in the knowledge of the Hubble constant; (2) the value of $\Omega = \rho/\rho_{\mathrm{cr}} \leqslant 1$ is uncertain, we assume that $\Omega \approx 1$, although this needs to be checked; (3) the nature of the hidden mass remains unknown (most likely, it is mainly of nonbaryonic nature, and hence is not essential here).

Certainly there should also exist some lepton asymmetry, although no experimental data about it are available, since most of the leptons and antileptons exist in the form of the up to now unobservable relic neutrinos and antineutrinos. Within a wide class of theories the relation $n_\mathrm{B} = n_\mathrm{L}$ holds true; however, some scenarios exist wherein n_L and n_B have different signs and magnitudes.

Numerous attempts are known to solve the problem of BAU *without abandoning the conservation law of the baryonic number* (items 1,2).

1. It is assumed (Alfven [1] and other authors) that at an early stage there existed a plasma that was primarily baryonic-neutral (hot Universe with a temperature $T \gg M_\mathrm{B}$) but later underwent the spatial separation of the baryonic charges

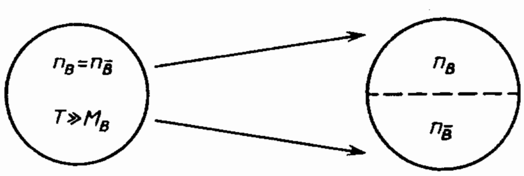


FIG. 1.

However, nobody succeeded in inventing a sufficiently effective separation mechanism. A shadow of hope is provided by a hypothesis concerning the role which may be played by the primordial black holes (Hawking [2], Zel'dovich *et al.* [3])

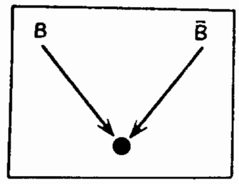


FIG. 2.

If, due to some reasons, the capture of antibaryons is more intensive, the excess baryonic charge arises outside the black holes.

2. Another hypothesis admits that the primordial matter was cold and contained *positive* baryonic charge (baryons or, more likely quarks). The primary entropy was either $S = 0$, or $S \sim B$ (S is the entropy, B the number of baryons in some region of the Universe). It grows in the irreversible processes in the course of expansion of the Universe, and the entropy increases to a value $\sim 10^9$.

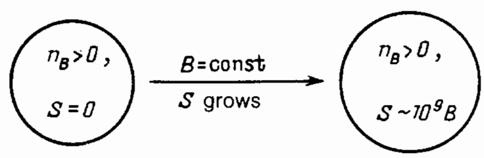


FIG. 3.

Hypotheses of this sort are *completely ruled out* within the scenarios of an inflationary Universe.

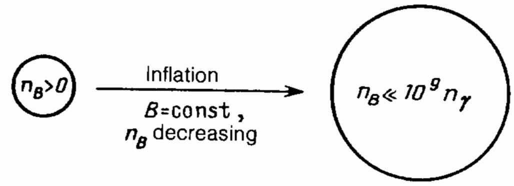


FIG. 4.

3. The third approach to the BAU problem *abandons the baryonic charge conservation* (S. Weinberg, 1964 [4]; A. D. Sakharov, 1967 [5]; V. A. Kuz'min, 1970 [6]).

Three basic conditions for cosmological formation of baryonic asymmetry

I. Absence of baryonic charge conservation.

II. Difference between particles and antiparticles, manifesting itself in the violation of CP-invariance.

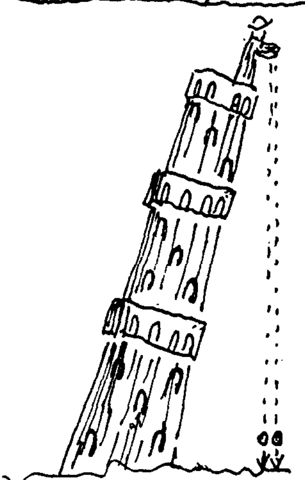
III. Nonstationarity. Formation of BA is only possible under nonstationary conditions in the absence of local thermodynamic equilibrium.

Let us discuss these conditions.

I. In the sixties the proton lifetime was estimated to be $\tau_{\text{prot}} > 10^{29}$ years. Now $\tau_{\text{prot}} > 10^{31} - 10^{32}$ years. In the sixties the conviction that the baryon charge conservation is an absolutely exact law was dominating.

Nevertheless, it was also admitted in the literature (Lee and Yang [7]), that this conclusion is not strictly obligatory, since no gauge field exists corresponding to baryon charge.

Nowadays two <u>exact</u> conservation laws are Known. These are the energy and electric conservations. In the both cases <u>a long-range</u> arises around the carrier of the conserved quantity. This is: the field of acceleration $a = \frac{GM}{r^2}$ in the case of the energy conservation Energy = mass $M \times c^2$; the electric field $E = \frac{q}{r^2}$ in the case of electric charge. This field are guarantors for the energy (mass) and charge conservation. Most precise experiments made to check the equivalence of the inertial and gravitational masses did not reveal any other long-range forces. The prototype is the <u>Galileo Galilei</u> experiment!



Newton, Eötvos A.
Dicke R.H. and collaborator
Braginsky V.B. and Panov V.I

The up-to-date precision makes $10^{-13} - 10^{-14}$

Once the baryon-charge-to-mass ratio variation makes $\sim 10^{-3}$, it is concluded that the baryon force should make less than 10^{-10} of the force of gravitation.

Thus No reason to expect that the baryon number conservation is equaly fundamental law as the other two! Indeed, it is violated in the GUT!

FIG. 5.

In the modern theories of "Great Unification" the nonconservation of baryonic and leptonic charges is the natural consequence of the quarks and leptons unification in one multiplet. The first theory of this sort was proposed in 1973 by Pati and Salam (integrally charged quarks). In 1974 Georgi and Glashow proposed their minimum SU(5) GUT that retains its importance up to now; there are leptoquark bosons X and Y with the charges $\pm 4/3$, $\pm 1/3$. Their two decay modes are

$$X_{4/3} \rightarrow \begin{cases} U_{2/3} + U_{2/3}, & (1) \\ e_+ + \bar{d}_{1/3}, & (2) \end{cases}$$

or, reading (1) from right to left,

$$\begin{cases} U_{2/3} + U_{2/3} \rightarrow X_{4/3} \rightarrow e_+ + \bar{d}_{1/3}, \\ B = 2/3 \qquad\qquad B = -1/3, \\ L = 0 \qquad\qquad\quad L = -1 \end{cases} \tag{3}$$

($\Delta B = \Delta L = -1$). The proton decay is governed by diagrams like this

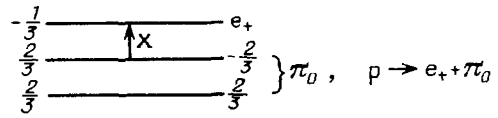


FIG. 6.

The search for this decay is as yet in vain, and the most convincing proof of baryon number nonconservation remains up to now the baryon asymmetry of the Universe.

Another mechanism (other than GUT) of the B-number nonconservation was suggested by G.t'Hooft, 1976 [9] (the case $T=0$) and by Dimopoulos, Susskind, 1978 [10], Linde, 1981 [11], Klinkhamer, Manton, 1984 [12], Kuz'min, Rubakov, Shaposhnikov, 1985 [13] (the case $T\neq 0$, $T\sim 10^2-10^4$ GeV).

At $T=0$ this is the quantum tunneling between different vacuum states of gauge fields. Its probability is

$$\sim e^{-4\hbar/\alpha}\ (\alpha = g^2/4\hbar = (137\sin^2\theta_W)^{-1}).$$

This effect is nonperturbative. Every term of the power series expansion of the function e^{-1/x^2} is zero.

At $T\sim 10^2-10^4$ GeV the baryon and lepton number non-conservation is strong!

II. CP-invariance violation was discovered in 1964 ($K_L^0\to\pi_+ + \pi_-$). I remember to have read in 1965 S. Okubo's paper[2)] where he stated that, in principle, CP noninvariance may lead to inequality of partial widths for particles and antiparticles in multichannel processes; but the total widths are the same. An example (experimental) of it may be the three-particle decay of K_L^0

$$K_L^0 \to \begin{cases} e_+ + \pi_- + \nu, \\ e_- + \pi_+ + \bar{\nu}, \end{cases}$$

$$\Gamma(K_L^0\to e_+ + \pi_- + \nu) \neq \Gamma(K_L^0\to e_- + \pi_+ + \bar{\nu}).$$

Later on, in the copy of my 1967 article presented to E. L. Feinberg I wrote the following inscription (in Russian):

Iz effekta S. Okubo
Pri bol'shoĭ temperature
Dlya Vselennoĭ sshita shuba
Po eë kosoĭ figure

Literal translation:

Out of S. Okubo's effect
At high temperature
A furcoat is sewed for the Universe
Shaped for its crooked figure.

Theoretically, the CP-violation is always due to appearance of complex phases that characterize the relative amplitude of some states or interactions. E.g. the CP-violation in the decay of K_L^0 is associated with the fact that this state is a superposition of K_0 and $\bar{K}_0$ with the phase difference $(\pi-\varphi)$, where φ is small.

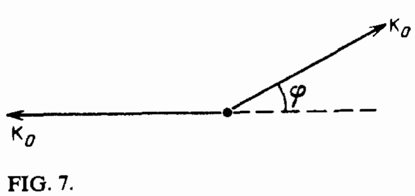


FIG. 7.

Three possibilities:

(A). The phase difference is built into the basic original equations of the theory.

(B). The phase differences result from a spontaneous symmetry breaking. They have a definite magnitude and random signs.

(C). The phase differences are random both in sign and magnitude.

Let us consider the possibility (B) and illustrate it by the following simplified toy-model with three scalar complex fields $\varphi_1, \varphi_2, \varphi_3$ with the potential:

$$U(\varphi_1,\varphi_2,\varphi_3) = \sum_i (\varphi_i\varphi_i^* - a^2)^2 - \gamma\sum_{i,j=12,23,31}|\varphi_i-\varphi_j|^2. \qquad (4)$$

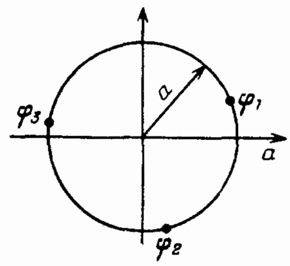


FIG. 8.

For $\gamma=0$ the phases of the complex fields φ_i are arbitrary. With $\gamma>0$ U depends on the phase differences. The T-symmetric configuration of Fig. 9a leads to an unstable equilibrium; the broken-T-symmetry state of Fig. 9b leads to a stable equilibrium with the phase angles differing by 120°. The points $\varphi_1, \varphi_2, \varphi_3$ can be ordered either clockwise or counterclockwise, modelling different sign of the T - (and CP) violations.

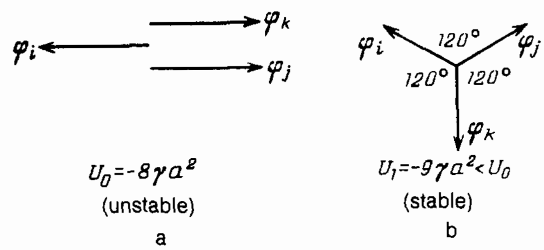


FIG. 9.

In case (B) domains can appear in the Universe that are characterized by different signs of CP-violation and hence by different signs of the baryonic and leptonic asymmetry (Fig. 10).

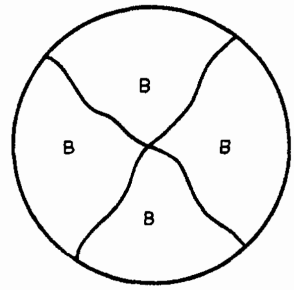


FIG. 10.

Within an inflationary model the domains will be colossal if the sign had been "chosen" before the inflation was over. In case the Universe is closed it can consist of only a single domain!

III. The necessity of nonstationarity. The equality of masses of particles and antiparticles implies that in thermodynamical equilibrium (where the entropy is a maximum) the number of particles and antiparticles has to be the same. A small difference in the numbers of particles and antiparticles results in a *negative* quadratic addition to the entropy. For massless free particles one has

$$\Delta S = -\frac{9}{2VT^2}\{2\Sigma(f-\bar{f})^2 + \Sigma(b-\bar{b})^2\}\ldots \qquad (5)$$

where f is the number of fermions of any type inside the volume V; b is that of bosons; the summation is carried out over the types of particles; T is the temperature; $VT^3 \approx \text{const}$;

$$B = \frac{1}{3}(q-\bar{q}), \quad L = \Sigma(l-\bar{l}).$$

q are quarks, and l leptons. If, initially, $q=\bar{q}$, $l=\bar{l}$, $b=\bar{b}$, the baryon and lepton asymmetry can arise only under nonstationary conditions when the substance is not in the state of thermal equilibrium (since the entropy cannot decrease). Expansion of the Universe leads to a nonstationarity which is greater when the Hubble constant $H=\dot{a}/a$ is greater. If the expansion is isotropic, the violation of the thermal equilibrium is due to the particle masses and to the density and temperature dependence of effective coupling constants.

During the years that passed after the first papers on BAU several different mechanisms were proposed for generating the baryon assymmetry (within the above three conditions). In the seventies most of the papers were devoted to the possibilities related to GUT (A. Y. Ignatiev, N. V. Krasnikov, V. A. Kuzmin, A. N. Tavkhelidze 1977 [14], M. Yoshimura 1978 [15], S. Weinberg 1979 [16], and many others). These works were based on the decay of leptoquark vector and Higgs bosons. It is necessary that the decay probability $1/\tau$ obey the inequality:

$$\frac{1}{\tau} \lesssim H = \frac{\dot{a}}{a}.$$

Multichannal decay modes like (1), (2) are considered and the difference in the particle and antiparticle widths due to CP-violation is calculated. It is true that the values of BAU close to observed ones are thus obtained. Still the GUT-based mechanism looks *doubtful* for two reasons:

1) First it does not blend well with the inflation scenario. The needed temperatures are 10^{15} GeV or even greater, whereas the reasonable inflation theories are not able to provide such a reheating temperature!

2) The second obstacle is in the low temperature violation of B and L conservation. This mechanism makes the substance approach equilibrium at $T = 10^2 - 10^4$ GeV.

The low-temperature process proceeds with the baronic and leptonic numbers decreased by the same quantity

$$\Delta B = \Delta L, \qquad B - L = \text{const}. \qquad (6)$$

If the asymmetry that has arisen at high (GUT) temperature is the same for B and L, it diminishes by many orders ($\sim 10^6$ or more) and practically disappears. This is just the case in the SU_5 models (Eq. (3)) and some other models.

Generally, there is no basis for thinking that the law (6) is absolutely exact. With an accuracy of 10^{-12} (greater than for baryon number) there are no long-range gauge fields corresponding to $B-L$. For a laboratory sample $B-L$ is simply the number of neutrons. In theories more sophisticated than the SU_5 of Georgi and Glashow the processes in which the boson doublets with the charges $(-2/3, +1/3)$ are involved conserve $B+L$ and not $B-L$.

$$d_{-1/3} + d_{-1/3} \to X_{-2/3} \to e_- + \bar{d}_{+1/3}$$
$$B = \frac{2}{3},\ L = 0 \qquad\qquad B = -\frac{1}{3},\ L = 1.$$

In some GUTs (SO(10), F(6)) such particles exist, and $B-L \neq \text{const}$, the neutron-antineutron oscillations are also possible in such theories (see Fig. 11)

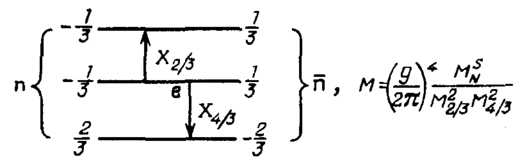


FIG. 11.

If the baryon-lepton asymmetry with $B \neq L$ arises at a temperature above the range $T = 10^2 - 10^4$ GeV (resulting from the decay of GUT bosons like $X_{2/3}$ or following another mechanism, e.g. via the squark and slepton decay, or via the cosmic strings decay), then a state will be established (with high precision) in the low temperature region which corresponds to the entropy maximum at given constant value of $B-L = \text{const}$ (and under the condition of electric neutrality),

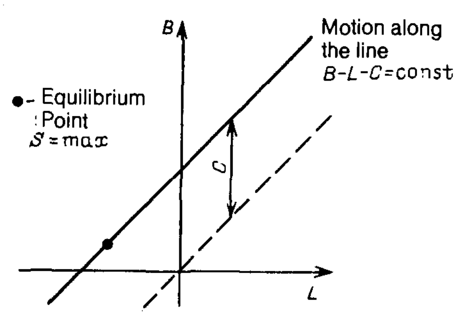


FIG. 12.

Then $\Delta S \sim -\{2\Sigma(q-\bar{q})^2 + 2\Sigma(l-\bar{l})^2 + \Sigma(b-\bar{b})^2\}$. In the maximum entropy state B and L have different signs: $B = 12/37C$, $L = -25/37C$ for example.

Shaposhnikov, Kuz'min and Rubakov consider also a possible role that may be played by the appearance of the baryon and lepton asymmetry in low temperature processes. It is not clear yet whether the asymmetry value about $10^{-9}n_\gamma$ can be attained in this way.

One of the currently most popular approaches is associated with the application of ideas inherent in supersymmetric versions of GUT (Affleck and Dine, [17], Linde [18]). Scalar fields with $B \neq 0$ and $L \neq 0$ are considered. For some linear combination of scalar fields the energy does not depend on the amplitude (or depends only in a nonperturbative way). Affleck and Dine suggest that it is these combinations that undergo especially large fluctuations in their phases and amplitudes (analogous property is exploited within the chaotic inflation hypothesis). The decay of these fields after the inflation or at the supersymmetry breaking temperature (assumed to be about $10^3 - 10^4$ GeV) results in the baryon and lepton asymmetry. Apparently the resulting asymmetry can be much greater than in the case of the GUT mechanism, and the fraction that survives in the low-temperature processes can also be greater since the temperature when the scalar particles decay may be of the order of temperature of low-temperature baryon-lepton non-conservation; the latter temperature may be larger in the presence of large baryon density (Ferrer and de la Incera [19]); it is not excluded that $B \neq L$.

The production of matter out of entropy provides entirely new possibilities in constructing cosmological models. They are effectuated, for instance, in various *inflationary* scenarios, as well as in oscillatory models.

Eternally oscillating model

The BAU problem is discussed in the framework of the evolutionary Universe. On the other hand, using the phenomenon of the matter production out of entropy, one can suggest a quasistationary oscillating model. Assumptions:

The cosmological constant is arbitrarily small and negative $\Lambda < 0$, $\varepsilon = \Lambda/8\pi G < 0$, $P = -\varepsilon > 0$.

The space geometry is *flat* on an average, and open.

The inequality $\varepsilon < 0$ may originate from the nonperturbative mixing of degenerate vacuum states. From $(\dot{a}/a)^2 = 8\pi/3G\,(\mathrm{const}/a^2 + \varepsilon)$, $\varepsilon < 0$ it follows that the pulsation period is finite

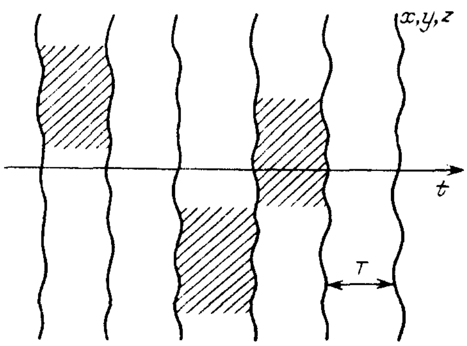


FIG. 13.

$$T = \frac{2\pi}{3}\left(\frac{8\pi}{3}G|\varepsilon|\right)^{-1/2},$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\left(\frac{\mathrm{const}}{a^2} + \varepsilon\right).$$

There is an infinite number of oases (dashed in Fig. 13), where the conditions do not significantly differ from what we observe!

Semieternal model

The same as before, but within a closed model (topologically S_3). 1) ε may be equal to zero. The maximum radius of the Universe in the n-th cycle ($\varepsilon = 0$) $R_n \sim T_n \sim S_n \to \infty$ at $n \to \infty$, where S_n is the total entropy of the Universe at the n-th cycle. 2) If $\varepsilon < 0$, then at large n the picture is quite like the eternal model

$$\varepsilon < 0; \quad T_n \to \mathrm{const}, \quad R_n \sim S_n^{1/3} \to \infty \quad \text{at} \quad n \to \infty.$$

In the S_3 semieternal model there exists the "origin of the world" $t = 0$! In this model one can also consider the evolution of the Universe at $t < 0$ by reflecting the arrow of time in the minimum entropy point. Assume that at the time moment $t = 0$ the Universe be in the false vacuum state. The entropy is then *minimum* ($S = 0$) at this moment and grows along both time directions. What can this formal possibility mean from the philosophical viewpoint, I do not know. An analogous picture, although singular at $t = 0$ ($a \sim \sin(kt)$) occurs in a closed model with negative space curvature. It is assumed that for $t = 0$ there exists the true vacuum with $\varepsilon < 0$. In the case of the T_3 space geometry (considered by Ya. Zel'dovich and A. Starobinski) the origin of the world is shifted to $t = -\infty$, the classical solution being exponential

$$a \sim e^{Ht}.$$

(The same as in the S_3-case, the false vacuum state is assumed); then such a model may be semieternal if in true vacuum $\varepsilon < 0$.

[1](The English text printed here is the re-edited version of the English translation appearing on pp. 65–80 of A. A. Friedmann: Centenary Volume, World Scientific, Singapore, 1990).

[2]Unfortunately I failed to find exact reference to this paper.

[1] H. Alfvén, Physics Today **24**, 28 (1971).
[2] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[3] I. D. Novikov, A. G. Polnarev, A. A. Starobinsky, and Ya. B. Zel'dovich, Astron. and Astrophys. **80**, 104 (1979).
[4] S. Weinberg, in "*Lectures on Particles and Fields*," edited by S. Deser and K. Ford (Prentice-Hall, Englewood Cliffs, N.J., 1964), p. 482.
[5] A. D. Sakharov, Pis'ma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).
[6] V. A. Kuz'min, Pis'ma Zh. Eksp. Teor. Fiz. **13**, 335 (1970).
[7] T. D. Lee and C. N. Yang, Phys. Rev. **98**, 1501 (1955).
[8] V. B. Braginskiĭ and V. I. Panov, JETP, **61**, 873 (1971)
[9] G. t'Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. **D14**, 3432 (1976).
[10] S. Dimopoulos and L. Susskind, Phys. Rev. **D18**, 4500 (1978).
[11] A. D. Linde, Phys. Lett. **100B**, 37 (1981).
[12] R. F. Klinkhamer and N. S. Manton, Phys. Rev. **D30**, 2212 (1984).
[13] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, Phys. Lett., **155B**, 36 (1985); *ibid.* **191B**, 171 (1987).
[14] A. Yu. Ignatiev, N. V. Krasnikov, V. A. Kuzmin, and A. N. Tavkhelidze, Izv. Int. Conf. on Neutrino, 1977 (In Russian) Nauka, M., 1978; Phys. Lett. **76B**, 436 (1978).
[15] M. Yoshimura, Phys. Rev. Lett. **41**, 281 (1978).
[16] S. Weinberg, Phys. Rev. Lett. **42**, 850 (1979).
[17] L. Affleck and M. Dine, Nucl. Phys. **B249**, 361 (1985).
[18] A. D. Linde, Phys. Lett. **160B**, 243 (1985).
[19] E. J. Ferrer and V. de la. Incera, Baryon asymmetry in high-dense postinflationary Universe, Moscow: Preprint FIAN N°223, 1988, p. 19.

А Сахаров