Viscosity effects in isotropic cosmologies

V. A. Belinskiĭ and I. M. Khalatnikov

L. D. Landau Institute of Theoretical Physics, USSR Academy of Sciences
(Submitted June 10, 1976)
Zh. Eksp. Teor. Fiz. 72, 3–17 (January 1977)

Some new types of evolution that arise in isotropic Friedmann cosmological models when allowance is made for bulk viscosity are described. The coefficient of viscosity is assumed to be a function of the energy density. For small and large values of its argument, this function is assumed to have a power-law behavior.

PACS numbers: 98.80.−k

1. INTRODUCTION

The role played by viscosity and the possible viscosity mechanisms in cosmology have already attracted the attention of many investigators (see, for example, [1–3] in which further references can be found), primarily in connection with attempts to explain the anomalously high entropy per baryon in the contemporary universe. Our present paper has a more general nature and aims at an approximate presentation of all possible types of cosmological solutions in Friedmann models when the influence of bulk viscosity is taken into account. It is assumed that the coefficient of this viscosity is a function of the energy density, which is approximated by arbitrary power-law dependences in the region of small and large values of its argument. Otherwise, the behavior of the coefficient of viscosity is hardly restricted at all.

It should be emphasized that new effects appear of course when the viscosity terms become of the same order as the other terms in the gravitational equations or even exceed them in order of magnitude. In these cases, the description of a dissipative process by means of just one coefficient of bulk viscosity may, strictly speaking, become invalid since none of the remaining terms in the expansion of the dissipative correction to the energy–momentum tensor in the velocity gradients may be small. However, allowance for the complete set (i.e., the transition from the hydrodynamic to the kinetic description) is extremely complicated and not amenable to simple analysis. Thus, the use of equations with hydrodynamic viscosity must be regarded as a model that describes dissipative processes, and it is naturally of interest to establish everything that such a simplified approach can give. However, in the light of what we have said it must be remembered that the results of this paper have, generally speaking, a qualitative nature and provide merely an idea of what could happen in cosmology when there is energy dissipation.

The most exotic effects occur in the stages of the cosmological evolution when the viscous terms in the equations become dominant. We shall call these stages "superviscous." One of these effects is the already noted[4] "matter creation" by the gravitational field at the initial time of the big bang (see the explanation below in the text to Fig. 4). It seems to us that this effect can be regarded as a phenomenological description of quantum particle creation in a strong gravitational field. The idea that particle creation could be interpreted in terms of classical viscosity has already been put forward by other authors.[5,6]

Some new effects also occur in the more realistic cases when the viscous terms are smaller but of the same order as the remaining terms in the gravitational equations. These effects include the unbounded accumulation of entropy during the later stages in the expansion of the universe (already noted in[4]), which occurs in certain types of solutions shown below in Fig. 2. In a closed model, these solutions describe expansion of space that is never replaced by contraction. Another phenomenon is the "slow big bang" described in the caption to Fig. 4. It occurs only in a closed model and represents a universe that begins from a singular state in the infinitely distant past ($t\to-\infty$). The matter begins to expand in this model infinitely slowly because the Hubble "constant" $H=\dot{R}/R$ at the initial time is zero despite the zero value of the radius R of the universe itself.

There also exist solutions in which viscosity does not lead to any qualitatively new effects.

2. GENERAL ANALYSIS OF EINSTEIN'S EQUATIONS

We write the Friedmann metric in the form[1)]

$$-ds^2=-dt^2+\frac{R^2(t)\,(dx^2+dy^2+dz^2)}{[1+k(x^2+y^2+z^2)/4]^2},\qquad (1)$$

where the cases $k=+1$, $k=-1$, and $k=0$ correspond to closed, open, and flat models.

In the case of isotropic cosmological evolution, there is no displacement of the matter layers with respect to one another, so that the shear (or first) viscosity does not appear, and the gravitational equations are the same as in the case of vanishing of the coefficient of this viscosity. Taking into account only bulk (or second) viscosity, whose coefficient we denote by ζ, we obtain the energy–momentum tensor in the form

$$T_i^k=(\varepsilon+p')u_iu^k+p'\delta_i^k,\qquad p'=p-\zeta u^k_{;k}.\qquad (2)$$

In homogeneous models, all scalar products depend only on the time t, and we can assume that the pressure p and the coefficient of viscosity ζ are functions of only the energy density ε. Taking a comoving co-

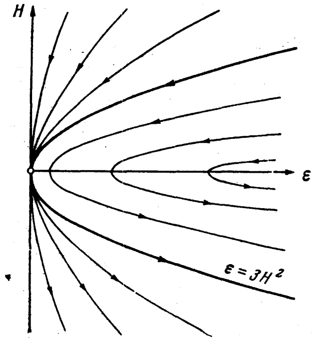

FIG. 1. Curves of ordinary Friedmann solutions when the coefficient of viscosity is zero. Above ($H > 0$) the expansion stage, below ($H < 0$) the contraction stage. The parabola $\varepsilon = 3H^2$ corresponds to the solutions for the flat model. Within it, the paths of the closed models, outside it, those of the open models. The figure corresponds to the case $\gamma < 4/3$. For $\gamma = 4/3$, all curves in the region of large ε run parallel to the parabola $\varepsilon = 3H^2$, but for $\gamma > 4/3$ they begin to converge on it. The equation of the phase curves is $\varepsilon - 3H^2 = 3k(\varepsilon/\varepsilon_0)^{2/3\gamma}$, where ε_0 is an arbitrary constant.

ordinate system in which $u^0 = 1$, $u^\alpha = 0$, we obtain for the effective pressure p'

$$p' = p - 3\zeta H, \tag{3}$$

where

$$H = (\ln R)^{\cdot} \tag{4}$$

is the Hubble constant (here and below the dot denotes derivatives with respect to t). The hydrodynamic equations $T^k_{i;k} = 0$ and the Einstein equations $R^k_i - \frac{1}{2}\delta^k_i R = T^k_i$ reduce now to the following three relations:

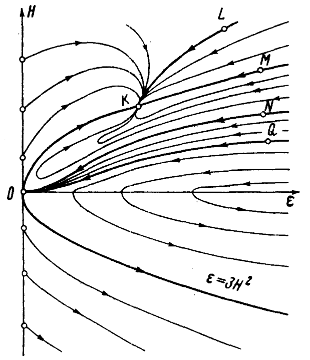

FIG. 2. Picture of the integral curves for the case $\nu < \frac{1}{2}$. The angle between the curves and the axis $\varepsilon = 0$ at the points of their intersections depends on the particular curve and the exponent ν. The figure corresponds to values $\nu < 0$. For $\nu = 0$, the angles are slightly different, and for $0 < \nu < \frac{1}{2}$ all curves touch the axis $\varepsilon = 0$, moving away from the points of tangency downward (as in Fig. 3 as well). The bundle of curves between the separatrices ON and OQ for $\nu = 0$ goes into the origin, forming now a nonzero angle with the horizontal axis; for $0 < \nu < \frac{1}{2}$, the angle is $\pi/2$. The disposition of the curves near the node K corresponds to the case when $3\gamma(1 - 2\nu) > 4$; if $3\gamma(1 - 2\nu) < 4$, then all the curves (except one) go into the node, touching the parabola $\varepsilon = 3H^2$. However, the qualitative behavior of the solutions described in the text does not depend on these details.

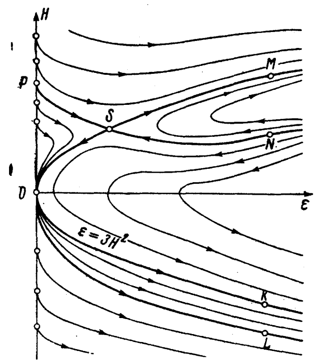

FIG. 3. Integral curves for $\frac{1}{2} < \nu < 1$. The picture corresponds to the case when $3\gamma(1 - \nu) < 1$. If $3\gamma(1 - \nu) > 1$, the bundle of curves between the separatrices OK and OL emanating from the origin O disappears. All curves below the arm OK of the parabola then begin from points of the axis $\varepsilon = 0$ with $H \neq 0$. Otherwise, the qualitative picture remains the same.

$$\dot{\varepsilon} = 3H(3\zeta H - w), \tag{5}$$

$$\dot{H} = {}^1/_3(\varepsilon - 3H^2) + {}^1/_2(3\zeta H - w), \tag{6}$$

$$\varepsilon - 3H^2 = 3kR^{-2}, \tag{7}$$

where w denotes the enthalpy, $w = \varepsilon + p$.

As usual, the hydrodynamic equation (5) is a consequence of Eqs. (6) and (7), and the relation (7) is a first integral of Eqs. (5) and (6). Equation (5) is related to the law of increase of entropy. In the case considered here, the entropy density σ is

$$\sigma = \mathrm{const}\cdot\exp\int d\varepsilon/w(\varepsilon), \tag{8}$$

and Eq. (5) can be written in the form

$$(\sigma R^3)^{\cdot} = 9\sigma R^3 \zeta H^2/w. \tag{9}$$

The quantity σR^3 describes the change with time of the entropy in a distinguished volume of comoving space (or in the whole of space for a closed model). All the factors on the right-hand side of Eq. (9) are positive by virtue of their physical meaning, from which it follows that the law of increase of entropy is satisfied for evolutions that develop in the direction of increasing time t. In what follows (in Figs. 1–6) the arrows on the integral curves indicate this physical direction of evolution.

We now note that for given equation of state $w(\varepsilon)$ and given dependence $\zeta(\varepsilon)$ Eqs. (5) and (6) describe a dynamical system in the phase plane (H, ε), the investigation of which can be conveniently made by qualitative methods. This system does not depend explicitly on the parameter k and applies equally to all three types of Friedmann model. At the same time, the integral (7) indicates the region of the phase plane in which the curves of each of the types lie. In the case of flat Friedmann model ($k = 0$), the integral curves lie on the parabola $\varepsilon = 3H^2$. Within this parabola, i.e., $\varepsilon > 3H^2$, we have the paths of the closed model corresponding to the case $k = 1$. Outside the parabola, where $\varepsilon < 3H^2$,

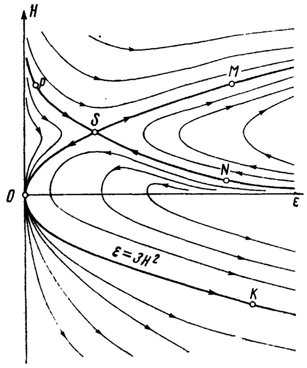

FIG. 4. Picture of integral curves for $\nu \geqslant 1$. To be specific, we have shown the case $\nu > 1$. If $\nu = 1$, the only change is that the separatrix NS and all the neighboring curves do not crowd toward the axis $H = 0$ as $\varepsilon \to \infty$ but toward the horizontal straight line $H = (3\gamma - 2)/9\alpha$.

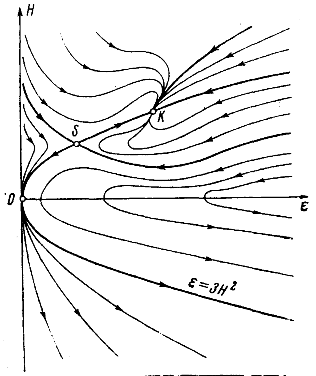

FIG. 6. Integral curves for the same asymptotic dependence of the coefficient of viscosity: for low values of the energy density $\zeta = \text{const}\cdot\varepsilon^{\nu_1}$, where $\nu_1 \geqslant 1$, while for large $\zeta = \text{const}\cdot\varepsilon^{\nu_2}$, where $\nu_2 < \frac{1}{2}$. In contrast to the preceding case (shown in Fig. 5) the curves $\zeta(\varepsilon)$ and $w(\varepsilon)/(3\varepsilon)^{1/2}$ intersect twice, which gives two singular points (H_0, ε_0)—the saddle S and the node K.

we have the paths of the open Friedmann model ($k = -1$).

The construction of the picture of the integral curves requires, above all, knowledge of the asymptotic behavior of the functions $\zeta(\varepsilon)$ and $w(\varepsilon)$ for large and small values of ε. It is interesting to note that, essentially, this is the only information needed for our qualitative analysis. In fact, besides this we only require that, between the points $\varepsilon = 0$ and $\varepsilon = \infty$, the functions $\zeta(\varepsilon)$ and $w(\varepsilon)$ do not have zeros or infinities and are sufficiently smooth. Under these conditions, as is readily seen from Eqs. (5) and (6), no integral curve can go off to infinity with respect to the variable H except either as $\varepsilon \to 0$ or as $\varepsilon \to \infty$. In other words, between these regions, i.e., for finite and nonzero values of ε, the curves can begin and end only at singular points with finite value of the coordinate H. It is easy to see that such points can lie only on the parabola $\varepsilon - 3H^2 = 0$ and only on its upper arm, where $H > 0$, and where the condition $3\zeta H - w = 0$ can be satisfied. It follows from this that the coordinates of these points (we shall denote them by H_0 and ε_0) can be found from the relations

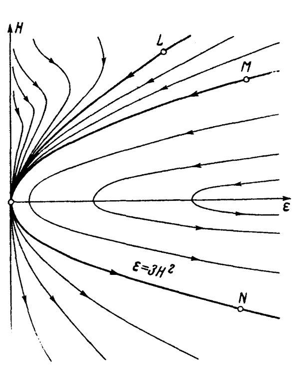

FIG. 5. Integral curves for the case when the coefficient of viscosity is approximated by a power-law dependence for both small and large ε, but the exponent in the limit $\varepsilon \to 0$ is greater than or equal to unity while in the limit $\varepsilon \to \infty$ the exponent is less than $\frac{1}{2}$. In addition, the curves $\zeta(\varepsilon)$ and $w(\varepsilon)/(3\varepsilon)^{1/2}$ do not intersect, so that there are no singular points (H_0, ε_0).

$$\zeta(\varepsilon_0) = w(\varepsilon_0)/(3\varepsilon_0)^{1/2}, \quad H_0 = (\varepsilon_0/3)^{1/2}. \tag{10}$$

Small variations of H and ε near H_0 and ε_0 can be expressed in the form of linear combinations of the two exponentials $\exp(\lambda_1 t)$ and $\exp(\lambda_2 t)$, with the characteristic numbers

$$\lambda_1 = 3\varepsilon_0\left[\frac{d}{d\varepsilon}\left(\zeta - \frac{w}{\sqrt{3\varepsilon}}\right)\right]_{\varepsilon=\varepsilon_0}, \quad \lambda_2 = -2H_0. \tag{11}$$

From the geometrical point of view, the position of the singularities (H_0, ε_0) is determined by the intersections of the curves $\zeta(\varepsilon)$ and $w(\varepsilon)/(3\varepsilon)^{1/2}$, and their nature (i.e., the sign of λ_1) is determined by the difference of the slopes of the tangents to these curves at their points of intersection. For $\lambda_1 < 0$, we have an absorbing, or attracting, node, and for $\lambda_1 > 0$ a saddle.[2)] If the singular point is a saddle, then one of the two separatrices which enter the saddle lies above the upper arm of the parabola $\varepsilon = 3H^2$ (in the region of an open model), while the other enters the singular point below this arm, from the region in which the curves of the closed model lie. The outgoing separatrices coincide with the part of the parabola $\varepsilon = 3H^2$ itself on which the saddle point lies, i.e., they are curves of the flat model.

If the asymptotic behavior of the functions $\zeta(\varepsilon)$ and $w(\varepsilon)$ at the origin and at infinity are given, then to find all the types of solutions one must consider all possible behaviors of these functions between the asymptotic regions and in each case establish the disposition, number, and nature of the singular points (H_0, ε_0). This analysis can be simplified by noting that the qualitative picture of the integral curves is determined by the number and nature of the singular points and does not depend on the remaining details of the actual dependence of the functions $\zeta(\varepsilon)$ and $w(\varepsilon)$. Therefore, without loss of generality we can represent these functions by any simple dependences, retaining only the most important properties.

We now consider more definite examples.

3. INVESTIGATION OF SOLUTIONS WITH CONCRETE DEPENDENCES $\zeta(\varepsilon)$ AND $w(\varepsilon)$

In this paper, we restrict ourselves to an equation of state of the form

$$w=\gamma\varepsilon, \quad 1\leq\gamma=\text{const}\leq 2. \tag{12}$$

With regard to the coefficient of viscosity, we shall assume that for both small and large values of ε the dependence $\zeta(\varepsilon)$ is a power law with corresponding power for each asymptotic region. In this section, we shall above all consider the special case when ζ is given by a single power function everywhere:

$$\zeta=\alpha\varepsilon^{\nu}, \quad \alpha, \nu=\text{const}. \tag{13}$$

We shall see that the results then obtained are determinative and enable one to construct the picture of integral curves for all possible more general cases (with power-law asymptotic behaviors) as well without any additional calculations.

Under the conditions (12) and (13), the curves $\zeta(\varepsilon)$ and $w(\varepsilon)/(3\varepsilon)^{1/2}$ have one and only one point of intersection if $\nu\neq\frac{1}{2}$.[3)] But if $\nu=\frac{1}{2}$, then these curves do not intersect. This last possibility corresponds to cases when there are no singular points. We shall not consider this possibility separately, but in the following section we shall consider more general variants of this kind. From (10) and (11) we obtain for the first characteristic number $\lambda_1=(\nu-\frac{1}{2})\gamma(3\varepsilon_0)^{1/2}$. Thus, for $\nu<\frac{1}{2}$ the resulting singular point (H_0, ε_0) is a absorbing node, while for $\nu>\frac{1}{2}$ it is a saddle.

We have investigated equations (5) and (6) qualitatively under the conditions (13) and (12) (for the case $\nu\neq\frac{1}{2}$) by compactifying the phase space of the dynamical system; for this the reader is referred to the Appendix. The resulting pictures of the integral curves in terms of the variables H and ε are shown in Figs. 2, 3, and 4. Figure 1 shows for comparison the paths of the standard Friedmann models without viscosity and with the equations of state (12).

Figure 2 shows the integral curves for the case $\nu<\frac{1}{2}$, when there is one node K on the upper arm of the parabola $\varepsilon=3H^2$. In the closed model (interior of the parabola) there are three classes of solutions, which we list in order of increasing importance of viscosity in them.

1. Solutions whose curves begin in the region $(H, \varepsilon)=(+\infty, +\infty)$, intersect the axis $H=0$, and go off into the region $(H, \varepsilon)=(-\infty, +\infty)$. These solutions describe one evolutionary cycle of a closed universe (from initial explosion to collapse) and are qualitatively the same as in the absence of viscosity. The first asymptotic terms in the expansion of these solutions for large ε are exactly the same as for a perfect fluid since the viscous term $3\zeta H$ in Eqs. (5) and (6) in the limit $\varepsilon\to\infty$ is negligibly small in the case considered. During the whole course of the evolution, the inequality $|3\zeta H|<w$ is maintained here, going over into $|3\zeta H|\ll w$ as $\varepsilon\to\infty$.

2. The class of solutions represented by curves of the bundle that come out of the region $(H, \varepsilon)=(+\infty, +\infty)$, pass through the origin $(H, \varepsilon)=(0, 0)$, and are bounded by the separatrices ON and OQ. These solutions describe an expanding closed Friedmann model in which a phase of contraction never occurs. The expansion begins with the usual initial big bang (with the same asymptotic behavior as without viscosity), but continues without end to $t=\infty$. During the final stages of infinite expansion, the viscous term $3\zeta H$ becomes decisive and of the same order as w, although, as before, the inequality $3\zeta H<w$ is everywhere satisfied. All curves of this family merge into a single curve as the origin O is approached, the equation of this curve near O having the form $3\zeta H/w=(3\gamma-2)/3\gamma$ (from which all the asymptotic behaviors are readily obtained). The radius $R(t)$ of the universe for these solutions increases monotonically from zero to infinity. The entropy of the universe (which is proportional to σR^3) increases from some initial value to infinity as well. These quantities increase as a power with the time t.

3. The third type of solution corresponds to curves that again come out of the region $(H, \varepsilon)=(+\infty, +\infty)$ but go into the node K (the curves between the separatrix ON and the upper arm of the parabola OM). They also describe an expanding closed model without contraction phase, but with an even stronger influence of viscosity. The start of evolution corresponds as before to a big bang, near which the viscosity is unimportant; however, on each curve there is now a point after which the inequality $3\zeta H<w$ is replaced by its opposite: $3\zeta H>w$ (at these points $3\zeta H=w$ and $dH/d\varepsilon=\infty$). In the superviscous phase that then follows, the expansion continues, as before, but the energy density ε begins to increase. The evolution terminates at the node K at $t=\infty$. The energy density and the Hubble constant tend to finite values H_0 and ε_0, while the radius $R(t)$ of the universe and the entropy, const $\cdot\sigma R^3$, become infinite. These quantities now tend to infinity as $t\to\infty$ in accordance with extremely rapid exponential laws with respect to the time t.

Outside the parabola $\varepsilon=3H^2$ in Fig. 2 we have the curves of an open Friedmann model. Here, we have the following three types of solution.

1. In the region $H<0$, below the lower arm of the parabola $\varepsilon=3H^2$ there are the curves of a contracting open universe. These curves begin at points lying on the axis $\varepsilon=0$ at finite H values. The angle between the curve and the axis $\varepsilon=0$ depends on the exponent ν and the particular curve (see the caption to Fig. 2). A common feature is however the fact that the contraction of the universe begins at a certain finite time t, with a certain finite value of the scale factor $R(t)$, and ends in the region $(H, \varepsilon)=(-\infty, +\infty)$ of the ordinary Friedmann singularity with negligible influence of viscosity. During the initial stage of contraction, viscosity is predominant and $|3\zeta H|\gg w$; during the final stage of collapse, $|3\zeta H|\ll w$. Evolutions with a beginning of this kind cannot have any sensible interpretation in the framework of classical hydrodynamics. Such curves require continuation into the region of negative values

of the energy density ε, and we shall not here discuss the possible meaning of such solutions. It is to be assumed that specification of the coefficient of viscosity in the form $\zeta = \alpha\varepsilon^{\nu}$ with $\nu < \frac{1}{2}$ is perfectly suitable for large values of ε but unrealistic for a classical fluid as $\varepsilon \to 0$. In this region one must have at least $\nu \geq 1$, since it is for such values of the exponent that, as we shall see later, the curves with the above exotic behavior disappear.

2. Above the upper arm of the parabola $\varepsilon = 3H^2$ there are the integral curves of an expanding open model. The bundle of curves that come out of infinity, $(H, \varepsilon) = (+\infty, +\infty)$, and is bounded by the separatrices KL and KM, goes into the node K. The solutions corresponding to the separatrices KL and KM themselves are independent types and will be noted separately. The remaining curves of this family describe evolution with an initial singularity of ordinary Friedmann type, near which the influence of viscosity can be ignored. During the initial stage $2\zeta H \ll w$, and thereafter $3\zeta H \leq w$ always, equality being attained only at the end point K, where the influence of viscosity becomes decisive. The evolution ends at K at $t = \infty$, and R and σR^3 become infinite.

3. The separatrix KL corresponds to a singular (and unique for $\nu < \frac{1}{2}$) solution which describes cosmological evolution of the same kind as in the preceding case but with strong influence of viscosity near the initial singularity $(H, \varepsilon) = (+\infty, +\infty)$ as well. For all curves between KL and KM the equation as $\varepsilon \to \infty$ is the same as without viscosity, i.e., $H = \text{const} \cdot \varepsilon^{1/2}$ (see the caption to Fig. 1), but the asymptotic equation of the curve KL as $\varepsilon \to \infty$ is $H = \text{const} \cdot \varepsilon^{1-\nu}$ (in the case considered, $1 - \nu > \frac{1}{2}$). This leads to quantitative changes in the initial behavior of the solution, but the expansion begins as before at a certain finite time with zero value of the radius R. The final behavior of the solution near the node K is in no way distinguished and is the same as for all the other solutions described in the preceding paragraph.

4. The last possible type of evolution in an open model corresponds to curves that begin at points of the axis $\varepsilon = 0$ with finite positive values of H and then go into the node K. The expansion begins in this case at a finite time t with a certain finite radius $R(t)$ and ends with $R(t)$ becoming infinite as $t \to \infty$ at the node K. On each curve there is a time at which the value of $3\zeta H - w$ becomes zero. This time separates the superviscous phase $(3\zeta H > w)$ from the final stage of evolution, in which the influence of viscosity is less $(3\zeta H \leq w)$ but nevertheless important. Solutions of this type, like those discussed earlier in the first paragraph (the curves below the lower arm of the parabola $\varepsilon = 3H^2$) have a beginning that cannot be realized in the framework of classical theory.

Finally, it remains to discuss the evolutions in the flat Friedmann model, whose curves lie on the parabola $\varepsilon = 3H^2$. Here there are possible solutions.

1. The lower arm of the parabola describes contraction ending with the usual singularity without appreciable influence of viscosity (with $|3\zeta H| \ll w$). However, at the start of the process viscosity plays a decisive role, $|3\zeta H| \gg w$, and, as a result, contraction begins with some finite value of the scale factor $R(t)$. The time corresponding to the start depends on the exponent ν. If $0 \leq \nu < \frac{1}{2}$, then the time is $-\infty$; but if $\nu < 0$, then the initial time is finite. In either case, the initial stage of this solution does not have a simple interpretation, as for all the cases noted earlier that begin at points of the axis $\varepsilon = 0$.

Out of the origin O there also comes the curve OK, corresponding to expansion in the flat model. The properties of the solution near the origin are exactly the same as in the preceding case. The evolution ends at the node K at $t = \infty$ with infinite values of $R(t)$ and the entropy. During the whole of the evolution, the energy density ε increases, despite the expansion, tending to a constant limit ε_0 as $t \to \infty$. This solution consists entirely of a single superviscous phase, for which $3\zeta H > w$. It is interesting to note that for $0 \leq \nu < \frac{1}{2}$ the expansion begins at $t = -\infty$ with some finite radius $R(t)$ and ends with $R(t)$ becoming infinite at $t = \infty$, i.e., we here have an example (of course, extremely special!) of a cosmological solution that is unbounded in time and free of singularities in the sense in which they are usually understood.[4)]

3. The part of the parabola MK describes expansion from a Friedmann singularity (without appreciable influence of viscosity) to an infinite value of the radius $R(t)$ and entropy equal to $\text{const} \cdot \sigma R^3$ at the node K as $t \to \infty$. The influence of viscosity is important during the late stages of the expansion, but $3\zeta H \leq w$ everywhere (equality only at K).

We now describe the new types of solutions that arise when $\nu > \frac{1}{2}$. Figure 3 shows the curves for the case $\frac{1}{2} < \nu < 1$. Compared with the preceding, this case is characterized by a more significant effect of viscosity in the region of large ε and a less significant effect in the region of low energy densities. Thus, we here already have classes of solutions whose asymptotic behavior near the origin $(H, \varepsilon) = (0, 0)$ is free of the influence of viscosity. But in the asymptotic region $\varepsilon \to \infty$ there is now no solution at all with negligible influence of viscosity. Let us consider first the types of evolution in a closed Friedmann universe, of which there are three.

1. Solutions describing a complete cycle of evolution in a closed model, whose curves go from the big bang region at $(H, \varepsilon) = (-\infty, +\infty)$ to the final singularity at $(H, \varepsilon) = (-\infty, +\infty)$, intersecting the axis $H = 0$. The initial singularity corresponds to a certain finite time t and a vanishing value of the radius R, but the asymptotic behavior of the solution depends on the coefficient of viscosity. The asymptotic equation of the integral curves in the region $(H, \varepsilon) = (+\infty, +\infty)$ is $3\zeta H/w = (3\gamma - 2)/3\gamma$, from which it follows that $3\zeta H$ and w are of the same order. During the final stage of evolution there is a superviscous phase, when $|3\zeta H| \gg w$, and the evolution ends (at a finite time) with a singularity of unusual form in which the radius R tends to some finite and non-

zero value but the derivative $\dot{R}$ becomes minus infinity. Features of this kind have already been encountered in in[4].

2. The class of solutions represented by curves covering the region NSM describes expansion of a closed universe without a subsequent contraction. The behavior of the solution near the initial singularity is exactly the same as in the preceding case. At the start of evolution $3\zeta H < w$, and the energy density decreases during the expansion, but then occurs a time after which $3\zeta H > w$ and ε begins to increase. The evolution ends with $3\zeta H \gg w$ and a singularity of the same unusual nature as in the preceding case: the radius of the universe R, increasing, reaches a finite value, and the derivative $\dot{R}$ tends to infinity (but now positive). This singularity corresponds to a finite time t.

3. A singular solution corresponds to the separatrix NS. The process begins here in the same way as the solutions described in the two previous paragraphs. The evolution consists of a single phase of expansion, during which $3\zeta H \leqslant w$ (equality at the point S) and ends at $t=\infty$ with R and the entropy const$\cdot\sigma R^3$ becoming infinite exponentially fast with respect to t at the saddle point S.

We now list the possible classes of solutions for an open model. As can be seen from Fig. 3, we here have five types of evolution.

1. Below the lower arm of the parabola $\varepsilon = 3H^2$ are the curves corresponding to contraction. The curves between OK and OL describe contraction of an open universe, beginning with ordinary Friedmann stages without viscosity. On these curves, near the origin O, the inequality $|3\zeta H| \ll w$ is satisfied, this being replaced at the end of the evolution by the opposite $|3\zeta H| \gg w$. The final singularity in the region $(H, \varepsilon) = (-\infty, +\infty)$ is of the same unusual type as in the solutions described above for the closed model that go into the region $(H, \varepsilon) = (-\infty, +\infty)$ (i.e., $t \to$ const, $R \to$ const, $\dot{R} \to -\infty$).

2. The curves below the separatrix OL correspond to solutions with a final singularity of the same nature as in the preceding case, but with an entirely different beginning. The process of contraction begins here with $|3\zeta H| \gg w$ at a certain finite time t with finite scale factor R. The curves begin at points of the axis $\varepsilon = 0$, touching it. Evolutions of this kind are incomplete and belong to the already considered (in the case $\nu < \frac{1}{2}$) class of solutions that require continuation into the region of negative energy densities, which in the framework of classical hydrodynamics is devoid of meaning.

3. All the paths above the upper arm of the parabola $\varepsilon = 3H^2$ have an initial stage of this unphysical kind, as can be seen from Fig. 3. The asymptotic behavior near the origin is here the same (except that now $H > 0$). All these solutions can be divided into three classes, depending on the nature of the final stage of the expansion. The paths below the separatrix PS give solutions that end above the unusual singularity, when $R \to$ const as $t \to$ const but $\dot{R} \to \infty$. During the whole of the expansion we have in this case $3\zeta H > w$.

4. The separatrix PS itself corresponds to expansion to an infinite value of R (at the same time, σR^3 also tends to infinity) as $t \to \infty$ in accordance with an exponential law. During the whole of the evolution $3\zeta H \geqslant w$.

5. Finally, the curves below the separatrix PS reach the origin O, near which the influence of viscosity disappears. At the start of evolution $3\zeta H \gg w$, but at the end $3\zeta H \ll w$, and the expansion during the final stage is purely Friedmann.

It remains to consider the solutions for the flat Friedmann model corresponding to the following three parts of the parabola $\varepsilon = 3H^2$.

1. The lower arm of the parabola OK corresponds to contraction from a purely Friedmann state (without significant influence of viscosity near the origin O) to a state with the previously discussed "unusual" singularity (R, $t \to$ const, $\dot{R} \to \infty$) characteristic of all the neighboring curves. At the start $|3\zeta H| \ll w$ and at the end $|3\zeta H| \gg w$.

2. The separatrix SO gives an evolution during the whole of which $3\zeta H \leqslant w$. The expansion begins at the point S as $t \to -\infty$ with zero value of the radius R and ends at the point O at $t = +\infty$ with infinite value of R. During the final stage of expansion $3\zeta H \ll w$, and there is no effect of viscosity. A solution of this type has been described by Murphy.[7]

3. The separatrix SM corresponds to a solution in which the expansion, which begins with zero value of R at $t = -\infty$, abruptly ends at a certain finite time with a finite value of R but infinite value of the derivative $\dot{R}$, i.e., with the unusual singularity discussed earlier.

Figure 4 shows the curves for the case when $\nu \geqslant 1$. The changes in this case are due to the even stronger influence of viscosity in the region of large ε and its weaker influence at low energy densities. There is no need to go through all types of evolution again; we mention only the main differences of the solutions corresponding to the curves of Fig. 4 from the preceding case (Fig. 3).

For a closed Friedmann model, the qualitative nature of the solutions and their asymptotic behavior change only near the initial time of the big bang. As in the preceding case, the asymptotic equation of the separatrix NS and all neighboring curves (as $\varepsilon \to \infty$) has the form $3\zeta H/w = (3\gamma - 2)/3\gamma$. It can be seen from this that in the region $\varepsilon \to \infty$ all these curves have a common horizontal asymptote $H =$ const for $\nu = 1$ and $H = 0$ for $\nu > 1$. In both cases the expansion begins at the time $t = -\infty$ with zero value of the radius R, but the Hubble constant H at the initial singularity is finite and even equal to zero in the general case when $\nu > 1$. An initial singularity of this kind can be called a slow big bang in the infinitely distant past. The final stages of evolution in these solutions, as can be seen from Fig. 4, can be divided into three possible classes, which are just the same as in the preceding case (Fig. 3). The final asymptotic behaviors of the solutions of all three types are the same.

In an open Friedmann model, the behavior of the integral terms is changed significantly only in the region of small ε. We see that for $\nu \geq 1$ the solutions finally disappear that we have called incomplete and which must be continued in some manner into the region of unphysical ε values. All integral curves corresponding to contraction now come out of the origin O, near which $|3\zeta H| \ll w$, and viscosity is not important. This initial stage is purely Friedmann. During the final stage of contraction $|3\zeta H| \gg w$, and we again have the same unusual asymptotic behavior as in the previous case. All curves describing expansion now begin in the region $(H, \varepsilon) = (+\infty, 0)$, where there is a true singularity. The expansion begins at a certain finite time t with zero value of the scale factor R. The energy density at the initial time is zero, increasing then with the later evolution. This phenomenon of matter creation by the initial singularity has already been noted by us in[4]. Near a start of evolution of this kind we have a superviscous phase, for which $3\zeta H > w$. As in the preceding case, the final stages of the evolution in these solutions can be divided into three types with the same asymptotic behaviors.

In the solutions for the flat model there are no qualitative changes from the preceding case at all.

4. CONSTRUCTION OF INTEGRAL CURVES IN MORE GENERAL CASES

The results of the preceding section enable us to construct readily the picture of the integral curves for more complicated behavior of the function $\zeta(\varepsilon)$ with power-law asymptotic behavior at the ends $\varepsilon \to 0$ and $\varepsilon \to \infty$. Indeed, the foregoing analysis contains a description of all possible pictures of the behavior of the paths at both small and large values of ε for any possible value of the exponent in the asymptotic dependence $\zeta(\varepsilon)$. Knowing the behavior of the curves in these asymptotic regions, we can readily join them up for each particular case determined by the nature of the singular points (H_0, ε_0), i.e., for each particular course of the function $\zeta(\varepsilon)$ between the points $\varepsilon = 0$ and $\varepsilon = \infty$.

As an example, let us consider the situation in which the coefficient of viscosity has the following asymptotic behavior:

$$\zeta = \mathrm{const}\cdot\varepsilon^{\nu_1}, \quad \nu_1 \geq 1 \quad (\text{as} \quad \varepsilon \to 0),$$
$$\zeta = \mathrm{const}\cdot\varepsilon^{\nu_2}, \quad \nu_2 < 1/2 \quad (\text{as} \quad \varepsilon \to \infty). \tag{14}$$

We choose this variant as the most reasonable from the physical point of view. In this case, in the region of small ε the curves do not begin at points of the axis $\varepsilon = 0$ (which occurs when $\nu_1 < 1$ and leads to a situation not readily interpretable, as we have already discussed), while in the region $\varepsilon \to \infty$ we do not have the "unusual" singularities mentioned earlier, for which the evolution ends with a finite and nonzero radius but infinite value of the derivative $\dot{R}$.

It is now easy to see that under the conditions (14) the curves $\zeta(\varepsilon)$ and $w(\varepsilon)/(3\varepsilon)^{1/2} = \gamma(\varepsilon/3)^{1/2}$ either do not intersect at all or have an even number of points of intersection. In the first case, there are no singular points (H_0, ε_0), and in the second there are an even number. The first of them (counting from the side of small ε) must necessarily be a saddle point, and then all the singular points follow in accordance with the rule saddle–node–saddle–node–... . We shall consider only the case when there are no singular points and when there are two. The cases when there are more than two can be omitted without loss of generality since they do not introduce qualitatively new types of solution.

The integral curves are now constructed by means of the results of the preceding section and simple considerations of a logical nature. Figure 5 shows the curves under the asymptotic conditions (14) when there are no singular points (H_0, ε_0). We shall not list once more all types of solutions and their asymptotic behavior, since they are all (except one) already described in Sec. 3. The new type of curve is represented by the bundle between the separatrices LO and MO. The only novelty is that this bundle now goes straight into the origin O. This corresponds to expansion in an open model from an initial singularity (without influence of viscosity) to an infinite value of R as $t \to \infty$ and, again, with negligibly small viscosity during the final stages of the expansion (near O).

Figure 6 shows the integral curves under the same conditions (14) when there are two singular points (H_0, ε_0)—a saddle S and node K. Using the information contained in Sec. 3, one can readily describe all the solutions corresponding to these curves.

APPENDIX

Here, in general features, we describe the qualitative investigation of the system of equations (5)–(6) in variables in which the phase space is compact. In such variables it is easy to establish the nature of all singular points of the system, including the infinitely distant ones, and the behavior of all the separatrices is uniquely determined. The inverse transformation enables one to establish then the picture of the behavior of the curves in the phase plane (H, ε).

We consider the case of a linear dependence of the enthalpy $w = \varepsilon + p$ on the energy density and an arbitrary power-law dependence of the coefficient of viscosity on ε:

$$\zeta = \alpha\varepsilon^{\nu}, \quad w = \gamma\varepsilon \quad (1 \leq \gamma \leq 2), \tag{A.1}$$

where α, ν, γ are constants. Instead of H and ε we introduce new variables φ and ψ by means of

$$\varepsilon/3H^2 = 1 + \mathrm{tg}\,\varphi, \quad 3\zeta H/\varepsilon = \gamma - 2/3 + \mathrm{tg}\,\psi, \tag{A.2}$$

and we replace t by the new time variable τ by means of the law

$$d\tau/dt = 3H/\cos\varphi\cos\psi. \tag{A.3}$$

With allowance for (A.1), the system of equations (5)–(6) now takes the form

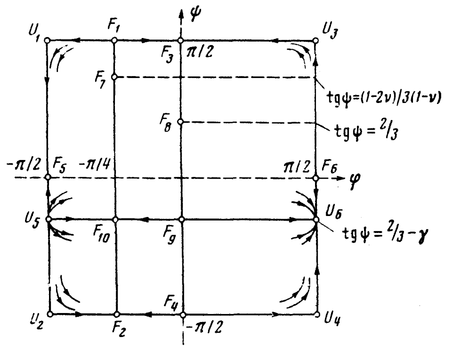

FIG. 7. Diagram of the singular points of the system (A.4). The position of the point F_7 depends on the value of the exponent ν and to be specific we show ones corresponding to values $\nu > 1$ (for $\nu = 1$ this point is absent). The arrows (in the direction of increasing value of the parameter τ) indicate only those directions that do not depend on the values of ν. The physical part of the square ($\varphi \geqslant -\pi/4$) below the straight line $\tan\psi = 2/3 - \gamma$ corresponds to positive values of H, and here growth of the parameter τ corresponds to increase of the physical time t. In the physical region below this straight line H is negative and the directions (of increasing τ) are opposite to the physical direction of evolution. This circumstance must be remembered when the picture is translated into the phase plane (H, ε).

$$d\varphi/d\tau = -(\sin\varphi + \cos\varphi)\sin\varphi\cos\varphi\sin\psi,$$

$$\frac{d\psi}{d\tau} = \frac{1}{2}\left[\left(\gamma - \frac{2}{3}\right)\cos\psi + \sin\psi\right]\sin\varphi\sin\psi\cos\psi \qquad \text{(A.4)}$$

$$+\frac{2\nu-1}{2}\left[\left(\gamma - \frac{2}{3}\right)\cos\psi + \sin\psi\right]\left(\sin\psi - \frac{2}{3}\cos\psi\right)\cos\varphi\cos\psi.$$

We shall investigate these equations only for $\nu \neq \frac{1}{2}$. The special case of the square-root dependence of ζ on ε must be considered separately since the transformation (A.2) is degenerate for $\nu = \frac{1}{2}$.

Positive values of the energy density ε correspond to variation of the variable φ in the interval $-\pi/4 \leqslant \varphi \leqslant \pi/2$, i.e., where $1 + \tan\varphi \geqslant 0$. The values of $\tan\psi$ are in no way restricted, and the variable ψ ranges over the interval $-\pi/2 \leqslant \psi \leqslant \pi/2$. However, the system (A.4), in contrast to the system (5)–(6), can also be continued smoothly into the region $\varphi < -\pi/4$, where $\varepsilon < 0$. Disregarding the physical meaning of this region, we shall consider formally the complete range of variation for both variables φ and ψ:

$$-\pi/2 \leqslant \varphi, \ \psi \leqslant \pi/2.$$

Thus, the phase space of the dynamical system (A.4) is a square whose edges (the straight lines $\varphi = \pm\pi/2$ and $\psi = \pm\pi/2$) are, as is readily seen, integral curves of of the system. Within the square there are three further integral straight lines: $\varphi = -\pi/4$, $\varphi = 0$ and $\tan\psi = 2/3 - \gamma$. All the singular points of the system (A.4) lie either on the edges of the square or on these three internal integral straight lines. In the general case there are 16 singular points, whose position together with all the straight integral curves (solid lines) are shown in Fig. 7. These straight trajectories are separatrices and divide the phase square into six rectangles, beyond the edges of which the integral curves can be continued only through singular points. Among the complete set of singular points there are six universal ones, in the sense that their behavior is completely independent of ν. These are the points U_1–U_6. The first four of them U_1–U_4 lie at the corners of the square and are saddles. The points U_5 and U_6 are noses. The behavior of the curves near these points are shown in Fig. 7. In addition, the directions on the rectilinear trajectories between the pairs of points F_1F_3, F_9F_{10}, and F_4F_2 are universal. All the remaining properties of the singular points depend on the intervals of variation of ν, and they must be considered separately for each case.

TABLE I.

F	$\nu < 1/2$	$1 > \nu > 1/2$, $3\gamma(1-\nu) > 1$	$1 > \nu > 1/2$, $3\gamma(1-\nu) < 1$	$\nu > 1$	$\nu = 1$	F	$\nu < 1/2$	$1 > \nu > 1/2$, $3\gamma(1-\nu) > 1$	$1 > \nu > 1/2$, $3\gamma(1-\nu) < 1$	$\nu > 1$	$\nu = 1$
F_1	RN	RN	RN	S	RN	F_6	AN–S	RN	RN	RN	RN
F_2	AN	AN	AN	S	S	F_7	S	AN	S	RN	No points
F_3	S	AN	AN	AN	AN	F_3	AN	S	S	S	S
F_4	S	RN	RN	RN	RN	F_9	RN	S	S	S	S
F_5	AN	RN–S	RN–S	RN–S	RN–S	F_{10}	S	S	AN	AN	AN

TABLE II.

F	φ_0	ψ_0	A	B_1	B_2
F_1	$-\pi/4$	$\pi/2$	$1/\sqrt{2}$	0	$(1-\nu)/\sqrt{2}$
F_2	$-\pi/4$	$-\pi/2$	$-1/\sqrt{2}$	0	$(\nu-1)/\sqrt{2}$
F_3	0	$\pi/2$	-1	0	$1/2-\nu$
F_4	0	$-\pi/2$	1	0	$\nu-1/2$
F_7	$-\frac{\pi}{4}$	$\mathrm{tg}\,\psi_0 = \frac{1-2\nu}{3(1-\nu)}$	$\frac{\sin\psi_0}{\sqrt{2}}$	$\frac{\sigma\sin\psi_0\cos^2\psi_0}{1-\nu}$	$-3\sigma\cos\psi_0$
F_8	0	$\mathrm{tg}\,\psi_0 = 2/3$	$-\sin\psi_0$	$1/2\gamma\sin\psi_0\cos^2\psi_0$	$(\nu-1/2)\gamma\cos\psi_0$
F_9	0	$\mathrm{tg}\,\psi_0 = 2/3-\gamma$	$-\sin\psi_0$	0	$(1/2-\nu)\gamma\cos\psi_0$
F_{10}	$-\pi/4$	$\mathrm{tg}\,\psi_0 = 2/3-\gamma$	$\sin\psi_0/\sqrt{2}$	0	$\sigma\cos\psi_0$

The nature of the ten singular points F is given in Table I, in which the following notation is adopted: RN, repulsive node; AN, attracting node; S, saddle. The abbreviations RN–S and AN–S denote complex equilibrium states: repulsive node–saddle and attracting node–saddle. These complex states arise only at the points F_5 and F_6, in the neighborhood of which Eqs. (A.4) cannot be linearized and require separate treatment. Near the remaining points F, the system (A.4) is linear. If we write

$$\varphi = \varphi_0 + \delta\varphi, \quad \psi = \psi_0 + \delta\psi, \qquad \text{(A.5)}$$

where (φ_0, ψ_0) are the coordinates of the singular points F, and $\delta\varphi$ and $\delta\psi$ are small deviations, then near all the singular points (except F_5 and F_6) the system (A.4) reduces to

$$d\delta\varphi/d\tau = A\delta\varphi, \quad d\delta\psi/d\tau = B_1\delta\varphi + B_2\delta\psi. \qquad \text{(A.6)}$$

The constants A, B_1, and B_2 together with the coordinates φ_0, ψ_0 are given in Table II, where we have introduced the notation $\sigma = [3\gamma(1-\nu) - 1]/3\sqrt{2}$. Note that for $\nu = 1$ the constants B_1 and B_2 for the points F_1 and F_2 are zero. To construct the integral curves, one must here make an additional investigation, but the nature of these points for the case $\nu = 1$ is given in Table I.

Equations (A.4) near the singular points F_5 and F_6 require a special investigation, which we shall not make here. We merely point out that for $\nu < \frac{1}{2}$ there exists a bundle of curves that enter the point F_6 from below and the left (from the third quadrant) and have near F_6 a common tangent, so that, taken as a whole, F_6 is in this case a complicated node–saddle type state. But if $\nu > \frac{1}{2}$, the point F_6 only repels the curves. The main bundle in this case comes out of F_6 upward and to the left (into the second quadrant), having a common tangent. The point F_5 for $\nu < \frac{1}{2}$ attracts curves from the first quadrant (upward and to the right), and for $\nu > \frac{1}{2}$ there is a bundle of curves that go out of it downward and to the right (into the fourth quadrant). In the last case, F_5 is a complicated node–saddle state.

Having elucidated the behavior of the integral curves of the system (A.4) in the physical region of the phase space ($\varphi \geqslant -\pi/4$) for every range of variation of ν, we can, using the transformations (A.2) and (A.3), then establish the picture of the integral curves in the variables (H, ε); this is shown in Figs. 2–4.

1)We use a system of units in which the velocity of light and Einstein's gravitational constant are equal to unity. The metric is described in the form $-ds^2 = g_{ik}\,dx^i\,dx^k$, where g_{ik} has the signature $(-+++)$. Latin indices take the values 0, 1, 2, 3 and Greek the values 1, 2, 3 and $(t, x, y, z) = (x^0, x^1, x^2, x^3)$.

2)We shall not consider the special case $\lambda_1 = 0$ of tangency of the curves, regarding this as too special and unjustified. This case is considered in detail in[4].

3)We recall that here we do not include possible intersections or tangency of these curves for $\varepsilon = 0$.

4)In Murphy's solution,[7] which corresponds to a flat Friedmann model with viscosity of the form $\zeta = \text{const}\cdot\varepsilon$ ($\nu = 1$), the scale factor $R(t)$ is nevertheless zero at the start of expansion (as $t \to -\infty$).

1S. Weinberg, Ap. J. **168**, 175 (1971).
2J. P. Nightingale, Ap. J. **185**, 105 (1973).
3Z. Klimek, Acta Astronomica **25**, 79 (1975).
4V. A. Belinskiĭ and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. 69, 401 (1975) [Sov. Phys. JETP 42, 205 (1976)].
5L. P. Grishchuk, Zh. Eksp. Teor. Fiz. 67, 825 (1974) [Sov. Phys. JETP 40, 409 (1975)].
6Ya. B. Zel'dovich and I. E. D. Novikov, Stroenie i Évolyutsiya Vselenoĭ (Structure and Evolution of the Universe), Nauka (1975).
7G. L. Murphy, Phys. Rev. D8, 4231 (1973).

Translated by Julian B. Barbour

A new type of radioactive decay: gravitational annihilation of baryons

Ya. B. Zel'dovich

Institute for Applied Mathematics, USSR Academy of Sciences, Moscow

(Submitted July 21, 1976)
Zh. Eksp. Teor. Fiz. 72, 18–21 (January 1977)

Some considerations are presented in favor of baryon-number nonconservation at the elementary particle level if the strong gravitational interaction at short distances is taken into account. A rough and unreliable estimate is given for the decay time of nuclei according to this mechanism.

PACS numbers: 23.90.+w, 12.25.+e, 11.30.Er

The application of gravitation theory to elementary particles leads to the conclusion that processes are possible which would manifest themselves as the conversion of baryons into neutral particles, e.g., $2N = (\pi^+, \pi^-, \pi^0)$ with apparent nonconservation of baryon number (baryonic charge, in the original). In these processes the rest mass of the baryons is totally converted into energy. Such a process may occur spontaneously in an atomic nucleus and even with an individual nucleon, $p \to e^+ + \pi^0$. Electrically neutral system can undergo oscillations, similar to the kaons; thus, the hydrogen atom can go over into its antiatom: $H = (pe^-)' \rightleftarrows (\bar{p}e^+) = \bar{H}$. At extremely high temperature processes of the type $\nu + \nu \rightleftarrows N + N$ become possible.

Extremely rough estimates for nuclei (stable with respect to the known decay modes) yield a lifetime of the order of 10^{45} years, which does not contradict the experiments of Reines[1] and of others. There are small chances that the probability is substantially larger, but the opposite result is also possible after a consistent theory will be developed for this phenomenon. Experimental detection of the presumed decay type is extremely difficult. On the other hand, a high-temperature reaction is likely to become comparable to other processes near the cosmological singularity, when the characteristic time is of the order of the Planck time 10^{-43} s. We assume that no new fundamental lengths will appear between the experimentally studied region of lengths and times and the Planck units of length and time[2] and that no fundamental change of the laws of nature occurs there.

What is the basis of this hypothesis? It has been known for a long time that the mass of a three-dimensionally closed universe vanishes identically. The local