# Gravitation, charges, cosmology, and coherence

Ya. B. Zel'dovich

*Institute of Applied Mathematics, Academy of Sciences of the USSR, Moscow*

Usp. Fiz. Nauk **123**, 487–503 (November 1977)

PACS numbers: 98.80.Dr, 97.60.Lf, 95.30.Cq, 95.30.Sf

## CONTENTS

## 1. INTRODUCTION

In recent years, especially in connection with the theoretical discovery of the quantum evaporation of black holes (Hawking, 1974[1]) (there is splendid review in the Uspekhi by Frolov[2]), exceptional interest has developed in the fundamental changes in the physical laws when strong gravitational fields must be taken into account. It has been stated that baryon charge is not conserved in black holes and that fundamental principles of quantum mechanics are violated, namely, the deterministic nature, i.e., the predictability, of the wave function (Hawking, 1975, 1976[3,4]). By "nonconservation of baryon charge" two things are meant.

1) Suppose that the baryon charge is coupled to some vector meson field like electric charge is coupled to the electromagnetic field (the four-vector of the potential). Suppose further that the corresponding meson has a rest mass, in contrast to the photon. In flat space, in the absence of gravitation, the static potential of $N$ baryons satisfies the inequality

$$A_{4N} > Nq\frac{e^{-ar}}{r}, \tag{1}$$

where $q$ is the charge of one baryon. This inequality replaces the electrostatic expression $A_{4E} = Nq_e/r$, $E_E = Nq_e/r^2$. In the space surrounding a black hole—in the Schwarzschild metric—it is found that $A_{4N} \equiv 0$ for $r > r_g$ (this result is obtained from the condition that $A_{4N}$ be finite at $r = r_g$; $r_g$ is the gravitational "Schwarzschild" radius of the body), whereas for the massless electromagnetic field the Coulomb law $E = Q/r^2$ is unchanged. Therefore, the vector meson field does not enable one to determine the baryon charge inside the black hole. Initially, "nonconservation of baryon charge" was used in connection with this disappearance of the exterior meson field. Later, in connection with the evaporation of black holes, which we shall discuss in the next section, it was noted that disappearance of the exterior meson field is needed if a black hole is to evaporate completely; an electrically charged black hole has a metric that differs from the Schwarzschild metric (the Reissner–Nordström metric) and cannot evaporate completely until it has freed itself of the charge.

2) The second meaning associated with "nonconservation of baryon charge" comes from the theory of evaporation of black holes. A black hole loses mass in accordance with a definite law, $dM/dt = -k/M^2$. It follows from the foregoing that this law does not depend on the baryon charge of the black hole. At a definite stage of the evaporation, when the radiation temperature is of the order or higher than the proton rest energy, the black hole also emits baryons and antibaryons, equally in the first approximation.

The evaporation of a baryon is accompanied by an antibaryon falling into the black hole; the evaporation of an antibaryon, by a similar fate for a baryon. But, as we have seen, the baryon charge of the black hole remains on the average constant during the evaporation. Suppose a black hole was formed by the collapse of matter (in contrast to antimatter) and has a positive baryon charge. In this case, it retains this baryon charge on the average right up to end of the evaporation. But the end of the evaporation occurs abruptly, in the perfectly definite time $t = M^3/3k$, after which the black hole completely disappears from our space. In our space, there remains nothing but the dispersed evaporation products; after the evaporation of the black hole, there do not persist at the position it occupied any perturbations of the metric or any umbilical cord joining our space to that region within the black hole that contained the collapsed baryons. These baryons have completely disappeared from our space, and in this sense one can speak of the nonconservation of baryon charge in our space.

The fate of the baryons within a black hole remains an open question. In classical (nonquantum) gravitation theory, i.e., in the general theory of relativity, a true singularity must of necessity arise within the black hole. The behavior of baryons that fall into this true singularity is unknown. However, the disappearance of the baryons from our space is not related to the singularity. There is no need to give up continuity of the baryon world line, i.e., to give up the microscopic formulation of the baryon charge conservation law.

The observed nonconservation of baryons is due to the appearance during the collapse of an horizon (in the

0038-5670/77/2011-0945$01.10

simplest case, the Schwarzschild sphere $r=r_g$) and the evaporation of the black hole associated with this horizon.

Let us now trace the history of this problem and sketch the probable stages in the further development of the theory. The proof that a massive vector field vanishes outside a black hole is completely analogous to the proof in Ginzburg and Ozernoĭ's well known paper[5] (1964) that an external magnetic field does. In the investigation of the nonspherical components of the gravitational field metric of a black hole a general principle was formulated very early (Doroshkevich, Zel'dovich, Novikov, 1965[6]): All quantities that can be observed outside a black hole depend only on the conserved quantities that characterize the black hole, i.e., its mass, angular momentum, and electric charge. Recall that at that time quantum evaporation of a black hole had not yet been considered. Later, this assertion was developed by several people and formulated succintly by Wheeler in the words "A black hole has no hair." This formulation presupposes that during the collapse process the external field "forgets" all the features in the structure of the collapsing mass; only information about the conserved quantities remains. And the baryon charge was not included among these conserved quantities. But I must admit that in[6] we did not form this conclude that baryons are not conserved.[1]

The problem of black hole evaporation also has a prehistory. Retrospectively, one should mention the paper[7] (Zel'dovich, 1962), in which it was shown that any number of baryons with any entropy can be compressed in such a way that the rest mass of the resulting body is arbitrarily small—because the gravitational mass defect almost compensates the rest mass and the internal energy of the compressed matter. The significance of this result is that any body, from a particle of dust to a star or pulsar, was previously assumed to be in a state of stable equilibrium but is in reality metastable, or unstable against transformation into a new supercompressed state with simultaneous transformation of the excess mass into radiation (protons, $e^+e^-$, $\nu\bar{\nu}$, ... pairs) that escapes to infinity.

The prolonged, virtually stable existence of matter in the ordinary state is due to the energy barrier, which separates the ordinary state of matter from the supercompressed state. A detailed treatment shows that the barrier (the quantity $A$ in the expression $e^{-2A}$ for the probability of the spontaneous process) is the smaller, the smaller is the initial mass under consideration. However, the nonquantum theory of gravitation applies only for $M \gtrsim \sqrt{G}$ (system of units $\hbar=c=1$; $G$ is the gravitational constant; in this system of units, the Planck mass unit $10^{-5}$ g corresponds to the mass $M_{\mathrm{Pl}}=(\sqrt{G})^{-1}$). At the limit of applicability of the theory $A \sim (GM_p^2/hc)^{-5/6}$, which is more than sufficient for the practical stability of ordinary matter and even matter compressed in neutron stars. But in principle the metastability remains. This result (the last scientific result I was able to tell Lev Davidovich Landau shortly before the catastrophe) was obtained for a matter distribution that does not qualitatively change the topology of Minkowski space. Actually, one could go even further back in time. It has long been known that the mass of a closed world is identically zero (Landau, Lifshitz[8]). Moreover, the electric charge of a closed world is also identically zero, but its baryon charge need by no means be zero![2]

Thus, energetically it is possible for a group of baryons in our space to be transformed into a group of baryons in a closed space separated from us. And it follows from the law of conservation of energy (which is a constituent part of general relativity or, better, a conclusion of general relativity) that there must appear in our space particles, on the average neutral, which inherit the energy of the group of baryons which have disappeared from our space.

Thus, the possibility of apparent nonconservation of baryons and transformation of their mass into energy existed in general relativity many years ago. Hawking's immense service is to have found the actual process that realizes this possibility. The process takes place after matter has collapsed to form a black hole.

For this process it is important that a black hole is essentially nonstationary, i.e., it is important that the collapsing matter approaches the gravitational radius asymptotically; the distance between the edge of the matter and the gravitational radius $(r-r_g)$ is halved in a time of order $r_g/c$, and this is precisely the period of the oscillations of the protons and other particles that, according to Hawking, evaporate from the surface of the black hole. Hawking's paper appeared after a number of studies of particle creation in simpler, ho-

---

[1] So long as we are dealing with Euclidean space, the inequality sign in (1) is such that the field outside does not disappear, but there is no Gauss theorem. However, if the meson mass is nonzero, the transition to the Schwarzschild metric of space around the black hole changes the situation, and the exterior field disappears. The actual disappearance of the static exterior solution does not depend on the magnitude of the rest mass, but the time of disappearance increases with decreasing mass (Starobinskiĭ).[12] Thus, at any finite time the dependence of the exterior field on the mass is smooth and the discontinuity in the stationary solution at $m=0$ occurs only at $t=\infty$.

Note finally that a massless field interacting with the baryon change does not exist in nature. More precisely, one can say that such a field would violate the equivalence principle: Distant bodies would interact not only gravitationally in proportion to their mass, but also through the hypothetical field, which would be proportional to the baryon number. Experiments show that if such an interaction exists, then it is at least $10^{50}$ times weaker than the electromagnetic interaction.

[2] For the electric charge, the relation $\operatorname{div}\mathbf{E}=4\pi e(n_p-n_e)$ and the reduction of the volume integral to a surface integral lead to $(n_p-n_e)=0$ for a closed world. Similarly, the four-momentum and, therefore, the mass can be expressed as a surface integral, which in the case of a closed world can be collapsed to a point. The physical meaning of the zero mass was elucidated by a consideration of half-closed worlds (Zel'dovich Novikov[21]).

mogeneous fields; this direction was initiated[3] by Parker.[9] Besides evaporating particles, a black hole polarizes the vacuum. The importance of vacuum polarization for particle creation in a gravitational field was demonstrated in[10] by Zel'dovich and Pitaevskiĭ. In the case of a black hole, the polarization energy density is negative and increases in absolute magnitude as $\varepsilon = -\text{const}/(r - r_g)$.[11,12] One can say that, as they disperse, the evaporated particles leave behind an ever greater negative mass of the polarized vacuum which compensates the (reducing) initial mass of the collapsing body. Further development of the theory encounters two very different groups of questions and problems.

When the black hole has a radius characteristic of the strong interaction ($2 \cdot 10^{-14}$ cm $= \hbar/Mc$), problems related to the structure of the proton, neutron, and mesons arise. Must these particles be regarded as elementary? Is the quark structure of hadrons important? Do there exist massive mesons or gluons coupled to the baryons and the baryon charge, or do there exist only gluon fields coupled to the quark color? The gluon field probably permits gravitational annihilation of quarks only as colorless triplets, i.e., in the form of one or several baryons. Do the gluons have mass and is it a bare mass or does the mass itself depend on some interaction (Higgs), and in such a case does the conclusion remain true that a massive field is pulled into the black hole by gravitation and disappears for $r > r_g$? At a somewhat smaller radius of order $10^{-17}$ cm, the energy of the emitting particles is of order of a few TeV (1 TeV $= 10^{12}$ eV). Here, the differences between the strong and the weak interaction disappear, as was pointed out long ago by Markov.[13] One can therefore expect nonconservation of parity and breaking of charge symmetry of particles and antiparticles.

In particular, a black hole could emit slightly more baryons than antibaryons, which it would preferentially absorb. This possibility was pointed out by Hawking,[3,4] and a particular concrete model was then considered by Zel'dovich.[14] In principle, one could in this way explain the observed charge asymmetry in the Universe and find theoretically the ratio of the baryon number density to the photon number density, which is a quantity that characterizes the present state of the Universe and has the value $\sim 10^{-8}$–$10^{-9}$.

At the radius $r_g = 2 \cdot 10^{-14}$ cm the black hole mass is $10^{13}$ g; at $r_g = 10^{-17}$ cm, it is $6 \cdot 10^{10}$ g. Therefore, the principles of Hawking's theory are still fully applicable: The gravitational field and the metric can be described by classical general relativity and the mass of the individual evaporating particles is negligible compared with the mass of the black hole. The evaporation of a particle can be treated on the background of a given metric without allowing for the back reaction of an individual particle on the black hole mass, whose variation can be regarded as a continuous process.[4]

However, when we wish to consider the deliberate production (with an accelerator) or spontaneous creation of a black hole, it is natural to consider black holes with the smallest possible mass. Particularization of the final stages in the evaporation of any black hole also leads one to consider black holes of small mass.

Hawking's theory itself determines its own limit of applicability. Omitting all dimensionless factors, let us write down the condition that the mass $M_{ev}$ of an evaporating particle be equal to the mass $M_{BH}$ of the black hole itself:

$$M_{ev} = \frac{\hbar\omega}{c^2} = \frac{\hbar}{c^2}\,\frac{c}{r_g} = \frac{\hbar}{c^2}\,\frac{c \cdot c^2}{GM_{BH}}.$$

The condition $M_{ev} = M_{BH}$ gives $M_{BH} = \sqrt{hc/G} = M_{Pl} = 10^{-5}$ g, $r = r_{Pl} = 10^{-33}$ cm. The characteristic Planck mass is obtained. Thus, the limits of applicability of Hawking's theory, which is based on general relativity, coincide with the assumed limits of applicability of classical (not quantum) general relativity itself.

But one cannot stop here. It is vitally important (in particular, for cosmology) to know what happens at smaller masses, and whether even black holes with smaller masses exist. Do they exist in nature and could they exist in principle?

There is no definite answer to any of these questions; there are only different opinions and differences of opinion.

Rather than enter into a discussion, let me list my suppositions.

1) A black hole with mass $10^{-5}$ g or less decays instantaneously, emitting, as a rule, a pair of particles. This process must be regarded as an individual quantum jump, on a par, for example, with the decay $\pi^0 \rightarrow 2\gamma$.

The characteristic time of the decay is of order of the Planck time, i.e., $\sim 10^{-43}$ sec, and therefore the mass uncertainty of the black hole is of the order of the mass itself.

2) The black hole, having emitted the particles, is transformed into a closed space, which is not connected to our space, i.e., the black hole completely disappears from our space.

3) The restriction on the quantitative application of Hawking's theory does not rule out decay. Stable formations with mass $10^{-5}$ g and radius $10^{-33}$ cm (or less) held together by gravitation do not exist in nature.

A future theory must give the decay time of unstable states with small mass.

In the region of very small masses, $M \ll M_{Pl}$, a com-

---

[3] There is even a first mention of this in Schrödinger's prewar paper.[31] For the cosmological consequences of particle production see[32–34].

[4] Note however that Ginzburg and Frolov[15] assume that already at a distance of order $10^{-17}$ cm fundamentally new physical phenomena could occur and all modern theories, including the theory of evaporation of black holes, prove to be inoperative.

pletely quantum language is needed. The lifetime of a black hole becomes so short that one must necessarily consider both processes—creation and evaporation—together.

Thus, if all the foregoing is true, one could for example, have the process

$$N + N = \mathrm{BH} = \nu + \nu + \mathrm{CW},$$

where CW is the abbreviation of Closed World, in this case a closed world containing two neutrons and two antineutrinos (i. e., a closed world with baryon charge $+2$ and lepton charge $-2$).

To avoid problems relating to meson fields, let us formulate the hypothesis more precisely.

We consider a world in which there exist only gravitation and massive Dirac particles with spin $\frac{1}{2}$. By hypothesis, these particles interact only gravitationally. This world resembles (perhaps, as a caricature of the original) the quantum electrodynamics of 1948–1960, which treated the system of Maxwell equations and the Dirac equation for electrons and positrons. The brilliant results obtained in this period are well known: shift of atomic levels, anomalous magnetic moment of the electron, and, to use the English expression, last but not least, the basic possibility of renormalization, i. e., the possibility of obtaining exact finite results by means of calculations in which divergent integrals occur in the intermediate stages.

One would like to carry out a similar program, replacing the electromagnetic field by the gravitational field. Hawking rightly emphasizes the difficulty of the gravitational problem compared with the electromagnetic (see also the earlier work of Feynman[16] and many others).

The constant of the electromagnetic interaction is dimensionless, $\alpha = e^2/\hbar c = 1/137$, and the perturbation theory series is expanded in ascending powers of $\alpha$, $\alpha^2, \ldots$. The constant $G$ of the gravitational interaction has the dimensions cm$^2$, or g$^{-2}$, or erg$^{-2}$ (for $\hbar = c = 1$) and therefore the power to which the maximal cutoff momentum occurs increases with increasing power of $G$, i.e., the divergence gets stronger.

A further complexity of gravitation compared with electrodynamics arises becasue the photon itself is not charged whereas the graviton has energy and momentum and is a source of gravitational field. Electrodynamics is effectively nonlinear only after the electron-positron field comes into play; gravitational theory is nonlinear by itself, even without a field of Dirac particles.

Hawking emphasizes that a solution of black hole type cannot be obtained in perturbation theory with a finite number of terms.

Finally, no conceivable gravitational experiments in the laboratory go beyond the linear nonquantum theory, and even in this framework a great deal remains undone, for example, the generation and detection of short gravitational waves.

For all these reasons, gravitational theory has advanced very much less than electrodynamics. What follows will be assertions, which may be probable and justified but they are not proved—in short, not results of a theory, but hypotheses.

The main one of these I shall call the zero-particle hypothesis. Namely, that there exist solutions describing two Dirac particles coupled together by gravitation to produce a system with zero mass. But the hypothesis is that one obtains a particle with $E \equiv p \equiv 0$, and not what one usually means by a massless particle, viz, $E \neq 0$, $p \neq 0$, $E^2 - p^2 = m_0^2 = 0$.

It is supposed that a zero-particle with $E \equiv p \equiv 0$ is in fact the natural quantum generalization of the concept of a closed world separated from our space. Can one say that a zero particle "exists" if it has neither energy or momentum? In the customary language of theory in which baryon conservation holds not all the quantum numbers of the particle are zero: its fermion charge is $+2$. We recall that in Dirac's theory free particles, i. e., ones not interacting with any other particles, are conserved trivially (the number of particles and the number of antiparticles are conserved separately).

If the gravitational interaction is added, then in the linear approximation already one can have creation and annihilation of particle-antiparticle pairs.

The Hawking theory of black hole evaporation directly suggests the concept of a zero-particle if the evaporation does not stop at any finite rest mass.[5)]

Can one assume in such a theory that the fermion charge—the difference between the number of particles and the number of antiparticles—is conserved? This conservation property holds for an individual vertex at which two fermion lines and a boson line of the gravitational field converge. But the same property holds for any finite set of vertices and for any real process in the gravitational field in any finite order of perturbation theory. From this there follows the law of conservation of the fermion charge and the concept of fermion charge as a state quantum number; in this system of concepts, one can speak of zero-particles with different quantum numbers.[6)]

However, if the zero-particle hypothesis is true, it can be interpreted differently. One can maintain that zero-particles do not exist in our space and then—for brevity—leave out the shibboleth "in our space." We shall assume that every other existent not in our space does not exist. Thus, zero-particles do not exist. In particular, we eschew the locutions " a zero-particle has fermion charge" or "fermions are in a zero-particle after the reaction." Instead of all this, we shall simply say that fermion charge is not conserved in a theory that includes the complete nonlinear theory of

---

[5)] In the following section devoted to cosmology, arguments in favor of complete decay will be given.

[6)] Odd baryon numbers are associated with nonzero half-integral spin and for this reason are probably inadmissible for a zero-particle.

gravitation, although conservation does hold in every term of the perturbation theory. Thus, one can have the reactions

$$f + f \rightleftarrows g + g,$$

where $f$ is a fermion and $g$ a graviton. Earlier, we would have said that $f+f$ can form a black hole, which evaporates, giving a neutral system—two gravitons.

Before, we would have said that at the same time zero-particle is formed; we now pass over that in silence.

## 2. COSMOLOGY AND BLACK HOLES

The whole problem of the quantum evaporation of black holes, the formation of zero-particles, and the nonconservation of fermion charge is tied up with the treatment of black holes of small radius and mass.

We recall that black holes first appeared in theoretical physics in connection with the study of massive stars. In 1939, Oppenheimer and Volkoff[17] considered neutron stars quantitatively in the framework of general relativity (albeit with simplifications); they found that there is a rigorous upper limit on the mass of cold matter that is capable of withstanding gravitation. Modern estimates give with greater confidence an upper limit of about two or three solar masses. Ordinary hot stars have masses right up to 30–50 solar masses. During their evolution, which proceeds rapidly for such stars in several million years, there may form in the center of such a star an iron core, in which nuclear reactions with release of heat come to an end. If the mass of the core exceeds the critical mass (which is only 10% or less of the mass of the complete star), it collapses, becoming a black hole. At a mass greater than the critical, there are no barriers anywhere between an ordinary density of a few grams per cubic centimeter to a density of order $10^{15}$–$10^{16}$ g/cm$^3$, which is characteristic for a black hole of mass $M \sim 2M_{\odot}$. However, the quantum emission of such a black hole is negligible: It emits like a body with an effective temperature of $10^{-7}$ °K. Therefore, the capture of matter and radiation from the surrounding medium exceeds by many orders the intrinsic emission of the black hole.

All the characteristic black hole evaporation phenomena occur with black holes whose mass does not exceed $10^{15}$–$10^{16}$ g (i.e., is less than $10^{-18}M_{\odot}$).

At the contemporary epoch, the formation of such black holes is precluded since to make such a body it is necessary to overcome an enormous energy barrier.

The practical possibility of forming light black holes is entirely restricted to the cosmological scenario of evolution of the Universe that envisages an infinite density of matter at the start of expansion near the singularity.

The idea of primordial black holes of cosmological origin was first put forward by Zel'dovich and Novikov.[18] They noted that at a high matter density black holes of low mass can form if the initial expansion law and the metric corresponding to it differ locally from the Friedmann solution.

On a large scale, it is well known that the Universe is uniform and expands equally in all directions.

But this does not prevent one assuming that perturbations—departures from the Friedmann solution—exist on a small scale. One can imagine a smooth, uncrumpled sheet of paper that is strewn with mountains and canyons if viewed through a microscope. The idea of small-scale inhomogeneities that are averaged out over large volumes appears very natural. Small regions with enhanced density and below average expansion velocity could be transformed into black holes. In order of magnitude, one can expect (with allowance for the relation $\rho = k_1/Gt^2$, where $k_1$ and the $k_2, k_3, \ldots$ introduced later are dimensionless factors of order unity) that at time $t$ after the onset of expansion matter occupying a sphere of radius $R = k_2ct$ and having mass $k_3\rho R^3 = k_4c^3t/G$ separates itself into a black hole with a Schwarzschild radius $R_g = 2GM/c^2 = k_5ct$. The formation of a black hole from matter within the horizon $R_h = ct$ entails initial perturbations of the metric of order unity on this scale. The process does not require a particularly strong change in the extent and density of the collapsing matter.

Note that in the early stage we are dealing with ultrarelativistic matter, so that one must assume in order of magnitude $p = k_6\varepsilon$, for example, $p = \varepsilon/3$, where $p$ is the pressure and $\varepsilon = \rho c^2$ is the energy density. At such a pressure, one requires an appreciable perturbation—of order unity—of the initial metric if a black hole is to form. It is known (Lifshitz, 1946[19]) that small perturbations of the metric in an ultrarelativistic gas give only small perturbations of the density by the time when $R \sim ct$, while for $R \ll ct$ the perturbations of the density and of the metric go over into a damped oscillatory regime.7) In cosmology too the formation of a black hole is a decidedly nonlinear effect. The existence of galaxies and clusters of galaxies indicate that there were certain deviations on the metric from the Friedmann solution on a large scale.8) According to an estimate of Novikov,[20] they are characterized by a dimensionless perturbation of order $10^{-2}$–$10^{-3}$ (for example, circumference $l$ equal to $2\pi(1 - 0.01)R$ or $2\pi(1 - 0.001)R$).

If a black hole with mass $10^{15}$ g is to be formed, the perturbation of the metric must be of order unity ($l = 2\pi \cdot 0.5r$ or $2\pi \cdot 0.3r$) on a scale $10^{17}$ times less than the scale of a cluster of galaxies.

In[18,21] Zel'dovich and Novikov noted an important feature of hypothetical primordial black holes. They form very early: for example, if $M = 10^{16}$ g then $R_g = 10^{-13}$ cm, so that $t = R_g/c = 3 \cdot 10^{-24}$ sec, and they form

---

7) Note that $R \sim \sqrt{t}$ for $p = \varepsilon/3$, and therefore $R/ct$ increases with the time.

8) In the theory of entropy perturbations, which lead to the formation of galaxies, initial perturbations of the metric near the singularity may be absent. For this and references to original investigations, see the monograph of[25] Zel'dovich and Novikov.

from matter of high density $\rho = 5 \cdot 10^{52}$ g/cm$^3$, which is ultrarelativistic. Matter which gets into a black hole does not expand after this. The surrounding matter that does not get caught in the black hole expands, the pressure doing work. Suppose that at the time of formation of black holes a definite fraction $\alpha$ of the matter goes into the black holes. while the fraction $1-\alpha$ remains in the normal state,[9] so that $\rho_{BH} = \alpha\rho = \alpha k_1/Gt^2$, $\rho_{NM} = (1-\alpha)k_1/Gt^2$.

After the radius of the Universe has increased by $n$ times, the mean density of black holes has decreased in inverse proportion to the volume by $n^3$ times, whereas the density of the surrounding matter has decreased by $n^4$ times. The relationship changes; instead of $\alpha:(1-\alpha)$ we obtain the density ratio $n\alpha:(1-\alpha)$. If 0.01 of the mass is transformed into black holes at the initial time, then after a hundredfold expansion these black holes already constitute 50% of the total mass; after an expansion by 10000 times, 99% of the total mass.

It was concluded on the basis of these arguments that in reality the Universe was also fairly smooth on a small scale—for otherwise the black holes formed at an early epoch would lead to the present time after expansion to an inadmissible matter density in the Universe.

The discovery of black hole evaporation has changed the situation. If the mass of the black holes is of order $10^{15}$ g, then, evaporating, they give rise to x and $\gamma$ rays. The sensitivity of instruments is such that these black holes could be detected at a density $10^8$ times less than the density at which one could detect passive black holes.

The discovery of evaporation strengthened the conclusion about the smoothness of the Universe's metric in the initial singular state on the scale $10^{-13}$ cm at the time $10^{-24}$ sec, which corresponds today (after expansion) to a scale of order 1000 km.

More subtle arguments enable one to rule out copious production of black holes with mass greater than $10^9$ g: evaporating, they would distort the spectrum of the fossil microwave radiation (2.7 °K) and interfere with nucleosynthesis.

On the other hand, when evaporation is taken into account one can allow copious production of the very smallest black holes with a mass from $10^5$ g (the quantum limit, the Planck mass) to, for example, ~1 g. Such black holes would have evaporated early, when the temperature was so high that thermodynamic equilibrium was completely established and "healed" all local perturbations due to the evaporation of the black holes. From the new point of view, considering the situation, we come to believe that copious production of black holes of the smallest possible mass, $\sim 10^{-5}$ g, is probable or even unavoidable since all fluctuations are of order unity at the corresponding characteristic time $10^{-43}$ sec.

If we were dealing with a few species of noninteracting particles, the energy of each particle at the time $t \sim 10^{-43}$ sec would be of the same order, $10^{-5}$ g, and black holes would be formed in an appreciable fraction of two-particle collisions.

Hawking's theory becomes necessary! Only the idea of evaporation of black holes can save us from the nightmare of microscopic black holes persisting to the present day and from the utopian requirement of absolute smoothness of the Universe on all scales right down to the quantum Planck scale $10^{-33}$ cm at the Planck density $10^{93}$ g/cm$^3$ of matter.

Thus, the only acceptable picture appears to be the formation and rapid evaporation of microscopic black holes immediately after the singular state. Physics leads to the same conclusion. But now cosmology has the word and imposes the answer to a question that physics cannot yet answer. From the point of view of cosmology, such a picture is possible and acceptable only if the evaporation of black holes *proceeds to the end*, to $M \equiv 0$. We have said above that the theory and Hawking's approximation (slowly varying classical gravitational field) become inapplicable when $M \leqslant 10^{-5}$ g.

If inapplicability of the theory meant that evaporation ceases, then today the Universe would be filled with particles that had not completely evaporated (maximons, planckeons). If this is not so, then evaporation does not stop at the point where Hawking's calculation becomes invalid. Perhaps the evaporation should now be regarded as a quantum jump (see above).

Thus, the conclusion of cosmology—the complete evaporation of black holes—favors the hypotheses advanced above. Thus, we have an indirect conformation that there can be a spontaneous process of transformation of baryons or a baryon into a lepton, for example, $P = \mathrm{BH} + e^+ = \mathrm{CW} + 2\gamma + e^+$ or $2N = \mathrm{BH} = \mathrm{CW} + 2\gamma$, where $P$ is a proton, $N$ a neutron, and CW, as before, a closed world that cannot be observed in our space. A first attempt to estimate the probability of such a process, without allowance for the quark structure of the baryon and gluon fields, was made in my note.[22] Even this estimate, which probably overestimates the probability, leads to a probability of annihilation of two neutrons in a nucleus of order $10^{-33}$ sec$^{-1}$ or less, which is virtually impossible to observe. Allowance for the quark structure of baryons will probably reduce this estimate further.

## 3. THERMODYNAMIC EQUILIBRIUM, COHERENCE, VACUUM

The evaporation of black holes makes it possible to pose new problems in thermodynamics. Imagine a vessel with heat impenetrable walls, i.e., a thermal bath filled with equilibrium radiation—electromagnetic, $e^+e^-$, and other pairs (if the temperature is sufficiently high), nuetrinos and antineutrinos, and gravitons. Note that it is not easy to make a wall impenetrable for neutrinos and gravitons.

At a given temperature of the bath, there will exist an equilibrium mass of a black hole such that the

---

[9] Below, we append the subscript NM to matter in the "normal" state; it is at a very high temperature but not in a black hole!

amount of energy absorbed in unit time is equal to the amount emitted. The connection between $T$ and $M$ is given by $T = \hbar c^3/8\pi GM$. Hawking[23] investigated the stability of this equilibrium. If the mass of the black hole is slightly increased, its temperature falls; at a constant *temperature* of the bath, there is a further increase in the mass of the black hole, i.e., the equilibrium is unstable. However, in the case of constant *volume* of a bath isolated from the external world, the increase in the mass of the black hole and decrease in its temperature are also accompanied by a decrease in the energy of the radiation in the bath and a fall in its temperature. If the bath has a small volume, the fall in the temperature of the bath is more rapid than the fall in the temperature of the black hole, the evaporation of the black hole restores the initial state, and the equilibrium is stable. Thus, at a given temperature, the volume of the bath also plays a role in the stability.

If there is an energy balance, i.e., if the energy absorbed is equal to the energy emitted, the spectrum and composition of the radiation (the relationship between the photons, neutrinos, and all the remaining particles) must not disturb the thermodynamic equilibrium in the bath. Otherwise, one could construct a perpetual motion machine of the second kind by extracting work from the nonequilibrium emission of the black hole. Hawking,[23] considering carefully the quantum evaporation process, proved the purely thermal nature of the radiation in all details, including the ratio of the probability for the emission of a particular number of photons of given frequency and polarization.

It should be noted that complete thermodynamic equilibrium is achieved only if the matter in the bath has a neutral composition and there are equal densities of baryons and antibaryons and of neutrinos and antineutrinos. This is because a black hole has no exterior fields, which would depend on the number of baryons or the number of neutrinos buried in the black hole.

Putting it differently, one can say that in the presence of a black hole the baryon and lepton conservation laws are violated, and therefore complete thermodynamic equilibrium is achieved in a symmetric system. Hawking pays particular attention to the thermodynamically equilibrium nature of the black hole emission. A nonequilibrium system containing a black hole increases its entropy and tends with the course of time to equilibrium, i.e., to the state that is most probable at the given total energy. Information about the initial state of the system is forgotten, i.e., lost.

In his well-known paper,[4] Hawking concludes that in the presence of black holes (or rather, when allowance is made for the basic possibility of their existence) the general logical scheme of quantum mechanics must be changed.

We recall that in quantum theory the equation for the wave function is unique and does not contain probability elements.

Probability enters through the answers obtained by measurements that include classical devices. The Schrödinger equation for an electron in an atom enables one to calculate $\psi(x, t_2)$ exactly if $\psi(x, t_1)$ is known but the question of whether an electron is in the volume $dv$ around a point $x_0$ is answered by a probability of the form $dP = |\psi(x_0, t)|^2 dv$.

The probability interpretation of quantum mechanics was accepted only with difficulty and struggle by physicists. But the point of view of Heisenberg, Born, and Bohr has passed the test of time and is the only one that is possible and correct.

In the future complete theory including the complete nonlinear and quantum gravitation, Hawking believes that the evolution of the wave function itself will become indeterminate and probablistic: He sees a manifestation of this indeterminacy in the evaporation of a black hole. Hawking emphasizes the difference between the picture he expects and the one generally accepted by the following comparison.

"The conclusion of this paper is that gravitation introduces a new level of uncertainty or randomness into physics over and above the uncertainty usually associated with quantum mechanics. Einstein was very unhappy about the unpredictability of quantum mechanics because he felt that "God does not play dice." However, the results given here indicate that "God not only plays dice, He sometimes throws the dice where they cannot be seen.""

Hawking's conclusion is very radical. But is not this conclusion connected with the fact that he considers a macroscopic black hole? The evaporation of such a black hole can be treated—as we have said above—semiclassically, with the quantized electromagnetic, neutrino, and other fields considered on the background of the classical metric of space and time, these tending asymptotically to the Schwarzschild solution. Could not this new and greater indeterminacy arise as a result of this macroscopic and semiclassical treatment of the situation? A particle that, say, consists of $10^{15}$ or $10^{18}$ carbon atoms absorbs any monochromatic coherent light and emits noncoherent thermal radiation corresponding to the temperature of the particle. But an individual carbon atom scatters the radiation coherently. Must one not treat the emission of a black hole at the quantum level? Can one not, and should one not formulate the theory with black holes in such a way that additional indeterminacy and incoherence do not arise?

We cannot answer this question confidently because there is no complete quantum gravitational theory. The considerations put forth below must be regarded as hypotheses. With this reservation, let us turn to the matter. For simplicity, we consider a theory in which there exist only the gravitational field and fermions. In this theory, there must exist black holes whose evaporation in the case $T > 2m$ ($m$ is the mass of the fermion) gives rise with virtually equal probability to fermions $F$, antifermions $\bar{F}$, and gravitons $g$. When the black hole reaches the Planck mass $m \lesssim \sqrt{G} = m_{\mathrm{Pl}}$, there is a quantum jump, $\mathrm{BH} \to 2F$, or $\mathrm{BH} \to 2\bar{F}$, or $\mathrm{BH} \to F + \bar{F}$, or $\mathrm{BH} \to 2g$. As the mass decreases, the number of particles into which the black hole must decay decreases,

TABLE I.

| Initial State \ Final State | $F+F$ | $F+\bar{F}$ | $\bar{F}+\bar{F}$ | $g+g$ |
|---|---|---|---|---|
| $F+F$ | $a_{11}$ | $a_{12}$ | $a_{13}$ | $a_{14}$ |
| $F+\bar{F}$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{24}$ |
| $\bar{F}+\bar{F}$ | $a_{31}$ | $a_{32}$ | $a_{33}$ | $a_{34}$ |
| $g+g$ | $a_{41}$ | $a_{42}$ | $a_{43}$ | $a_{44}$ |

and therefore we have restricted ourselves above to two-particle decays. By virtue of reversibility, the opposite processes of formation of black holes in two-particle collisions must occur:

$$2F \to \mathrm{BH},\quad 2\bar{F} \to \mathrm{BH},\quad F + \bar{F} \to \mathrm{BH},\quad g + g \to \mathrm{BH}.$$

Since black holes are unstable, we must consider processes of formation and decay of a black hole together. In this case, we obtain a matrix of 16 transition probabilities in Table I.

What distinguishes this table from any other $S$ matrix in the presence of several channels? A superficial difference is that the table includes processes that violate the fermion conservation law despite the fact that the original Lagrangian, and also the perturbation theory satisfy this conservation law.

In its turn, this superficial difference has a very deep internal origin: The formation of a black hole and its subsequent evaporation are connected with the separation from our space of some separate, topologically disconnected closed world. This closed world carries away fermion charge—positive or negative, and it is different for different channels. In classical theory, this closed world is nonstationary, and in it singularities must arise after a finite proper time of the closed world. We can draw conclusions about the properties of a quantum closed world only by analogy and by means of indirect considerations that generalize cosmological observations and theory. We retain in the quantum case the principal property of a closed world: the identical vanishing of its energy and momentum.

By what criterion are we to judge whether the formation of a closed world is a real event?

Among the processes listed in the table above, those that do no violate fermion conservation can also take place without the formation of black holes and closed worlds, for example, $F+\bar{F}\rightleftarrows g+g$, or all scatterings, for example, $F+F=F+F$. Now the question is this: Must we add the *amplitude* or the *probability* for two paths: 1) the direct process and 2) with the formation of a black hole and a closed world?

Hawking's suggestions correspond to the proposal to add the probabilities since for path 2) the state of the closed world which detaches itself from our space remains unknown and indeterminate. But one could perhaps develop a consistent theory in which the state of the closed world has no influence at all on what occurs in our space, i.e., a theory which one can reformulate in such a way that closed worlds are not mentioned explicitly in it at all. One of the consequences of this theory will be the calculation of the *amplitude* of processes that "in reality" take place through black holes and closed worlds with the concomitant addition to the amplitudes for the process with and without closed worlds. In such a theory, coherence will be violated only for processes having a macroscopic nature, and it will not be of such a fundamental nature.

Note that the decision not to consider closed worlds, is important for the formulation of the very concept of thermodynamic equilibrium. At a high temperature, one cannot rule out copious production of black holes and closed worlds as a result of thermal fluctuations of the density. The evaporation of black holes restores the energy balance, but must we not take into account the number of closed worlds, which varies with time, and the back reactions of the closed worlds (which "swim up to us" out of other dimensions—if we picture the four dimensional $x$, $y$, $z$, $t$ as imbedded in a space with larger number of dimensions) on us?

Fomin[24] has noted that the vacuum at zero temperature is unstable against the production of closed worlds; this has something in common with Wheeler's ideas,[29] according to which the metric of the vacuum contains not only zero-point vibrations of the transverse degrees of freedom—the gravitational waves—but also spontaneous variations of the topology such as the formation and disappearance of arms. At the same time it is clear that all these exotic processes in nature and in the future true theory will in a sense be eliminated or renormalized in such a way that the state of the vacuum at zero temperature is unique and the energy density $\varepsilon$ of the vacuum is identically equal to zero.

The last fact, i.e., $\varepsilon=0$, is a reliable consequence of physical and astronomical observations (the cosmological having the greater accuracy, $|\varepsilon|<10^{-8}$ erg/cm$^3$), along with the less certain assertion made above about the complete evaporation of black holes.

Observations are the more important the further from us (at present) is the ideal of a complete theory.

At the present time, observations do not refute the basic possibility of constructing a theory including gravitation with a single vacuum, without conservation of fermion charge but with retention of coherence and the general principles of quantum mechanics.

Considered technically and more narrowly, the possibility of constructing such a theory is related to the question of whether or not unitarity of the theory is violated when one decides to leave evaporated black holes out of account. The alternative, preferred, it seems, by Hawking (letter to the author, July, 1977) is a theory with infinitely degenerate vacuum containing on the average one black hole with Planck mass in each volume equal to the cube of the Planck length.[10)] The

---

[10)] Hawking developed this point of view in his talk at the Eighth Conference on General Relativity and Gravitation at Canada in August 1977.

requirement $\varepsilon = 0$ or $|\varepsilon| < 10^{-8}$ erg/cm$^3$, or, in Planck units $|\varepsilon| < 10^{-113}$, also remains true for the vacuum in this theory. To forestall misuse, let us say straight away that under no conditions can the vacuum become a source of energy.

Finally, let us point that the possibility of nonconservation of baryons and transitions between different states of the vacuum (instantons) is now also discussed in field theories in flat Minkowski space (Polyakov,[26] Belavin *et al*,[27] and t'Hooft[28]). It is possible that the methods developed by these authors will also be helpful in gravitation theory.

## 4. CONCLUDING REMARKS. GRAVITATION THEORY, ELEMENTARY PARTICLES, AND DIRAC'S LARGE NUMBERS

During the writing of this paper about black holes and nonconservation of baryon charge, it has become ever more clear to me that the subject of the paper is but part of a greater problem. The title of this last section must briefly characterize the questions at the heart of the problem. It could be that black holes are a special case of strong changes in the structure of space and time[29] that are important for the properties of elementary particles. The time approaches when we must put on the agenda the creation of a unified theory of all interactions, including gravitation of all particles and gravitons. The readers of this journal will hardly include people so naive as to expect the presentation in the following pages of a unified theory, least of all by a single author. But let me at least put forward some considerations about the possible form of this theory, and show how one can eliminate certain prejudices that have existed for more than 40 years.

The aim of unifying the theories of different phenomena is of course the most important principle of epistomology. Classical examples are the Newtonian theory of gravitation, which unified terrestrial attraction and celestial mechanics, and the Faraday–Maxwell unification of electricity and magnetism. During the last decade, successes have been achieved in the path toward unification of the weak interaction and electromagnetism. It is natural that one feels with ever greater urgency the need to unify the theories of elementary particles (including electrodynamics) and the theory of gravitation. Essentially, Einstein devoted half of his life to this problem—and did not achieve success. Simple dimensional arguments give an inkling of the possibilities—and of the difficulties.

The constants $c$, $\hbar$, and G (velocity of light, quantum of action, and the gravitational constant) have independent dimensions, and the number of these constants is such, that using them, one can express the units of length, time, and mass, and therefore, all the quantitative properties of all elementary particles.

For example, in a unified theory, the masses of the proton and the electron will be expressed in terms of $G$, $\hbar$, and $c$ with definite, theoretically determined numerical coefficients. Planck drew attention to this tempting possibility soon after his discovery of the quantization of light and the introduction of $h = 2\pi\hbar$.

After this there passed about 30 years without measurable success, and therefore—psychology has its own laws—increasing disenchantment.

Dirac,[30] considering this question, pointed out that the dimensionless quantities that should figure in a unified theory are very different from unity. For example, the proton mass is $m_p = 10^{-19} m_{\rm Pl} = 10^{-19}\sqrt{\hbar c/G}$. Putting it differently, one can say that the dimensionless ratio $Gm_p^2/\hbar c \approx 10^{-38}$. Dirac expressed the belief that such numbers cannot be obtained theoretically. In a theory, one usually introduces factors of 2, 3, $\pi$, and $e$ to different integral powers, positive and negative, but it would be very hard to take sufficient factors to give $10^{38}$ or $10^{-38}$. Dirac drew far reaching conclusions from his observation. We encounter large numbers in cosmology. For example, the age of the Universe, defined as $H^{-1}$, where $H$ is the Hubble constant, is equal to about $2 \cdot 10^{10}$ yrs $= 6 \cdot 10^{17}$ sec. But this is a dimensional quantity. Let us compare it with a quantity with the dimensions of time characteristic of a nucleon, say $\hbar/m_p c^2 = 6 \cdot 10^{-24}$ sec. Their ratio is $10^{41}$; if $10^{38}$ and $10^{41}$ differ by 1000 times, the exponents (i.e., the logarithms of the dimensionless numbers) 38 and 41 differ little, by less than 10%.

Dirac considered the number of nucleons in the complete Universe (assuming that it is closed) or in the observable part of the Universe, if it is open; this is of order $N = (c/H)^3 n$. If $n \approx 10^{-6}$ cm$^{-3}$, then $N$ is of order $10^{79}$, so that approximately $GM^2/\hbar c = N^{-1/2}$. Dirac concluded that these coincidences are not fortuitous! Written out in full, the first relation has the form $Gm_p^2/\hbar c = \hbar H/m_p c^2$, $Gm_p^3 c/h^2 H \sim 1$. The second relation amounts to a cosmological density which is of the order of the critical density:

$$\rho_c = nm = \frac{1}{6\pi G t^2} \sim \frac{H^2}{G}, \quad N \sim \frac{c^3}{H^3}\frac{H^2}{Gm}.$$

Comparing, we obtain $c^3/HGm = (\hbar c/gm^2)^2$, i.e., the same relation $Gm^3 c/\hbar^2 H \approx 1$. If Dirac's arguments were indisputable, the consequences would be immense. Formulas that combine cosmological quantities ($H$) and local quantities ($G, m$) herald the end of local physics and suggest some kind of lone-range effects: The mass of a (given) proton depends on the total number of protons in the Universe, i.e., other protons act on the (given) proton. In addition, in an evolving cosmology the Hubble constant decreases with the course of time, and the age of the Universe increases, so that $G$, $m$, $\hbar$, and $c$ cannot remain constant and at least one of them must change with the course of time.[11]

However, the recent years in the theory of elementary particles have posed new problems and brought new results which enable one to attack the problem of large numbers in a completely different way.

---

[11] One of the most recent theoretical investigations is[35]. Observations do not indicate the slighest hint of a change in these quantities.[36]

In a theory of interacting gauge vector meson fields one can find a combination of potentials to which there corresponds the absence of fields. Such a situation is well known in electromagnetism: if $A_\mu = \partial\varphi/\partial x^\mu$, then

$$F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu} = \frac{\partial^2 \varphi}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 \varphi}{\partial x^\nu \partial x^\mu} \equiv 0.$$

However, in the case of several interacting fields, different combinations $A_\mu^{(n)}$ (the subscript labels the coordinate and the superscript the number or color of the field) with vanishing fields describing the vacuum can be divided into classes between which there are no continuous transitions. Here there is no space to describe in detail and, more importantly, comprehensibly the resulting situation. Let us be content with saying that these combinations of potentials are distinguished by their topological structure in isotopic space. As the Romans said "Sapienti sat," which with a little licence we might render as "A word to the wise is sufficient" or, in a free translation of the Mulla Nasrudin: "Let those that know tell those that do not."

The transition from one combination of potentials without fields to another is possible only through a spacetime region with nonzero fields. This is a tunnel transition, which is impossible in classical theory but is possible in quantum field theory. Such a transition was considered by Polyakov and his colleagues[26,27]; later it was called an instanton. In contrast to a particle, which moves along a worldline, an instanton is a small region that is bounded not only in space but also in time. The probability for the appearance of an instanton in the vacuum contains the factor $\exp(-16\pi^2/g^2)$, where $g^2$, the charge of the field, is probably $\sim 0.17$, so that numerically we obtain $\exp(-10^3) = 10^{-430}$.

We have said all this here only to show that in modern theories one can now obtain numbers that differ very strongly from unity. More, one can formulate a law: In perturbation theory, one obtains factors of the type $2^n$, $\pi^k$, $g^m$, where the powers are determined by the order of the process. However, there exist processes that cannot be described at all by any order of perturbation theory. These include the spontaneous topology change associated with the tunnel transition. In this case, the natural numbers and the dimensionless charge are raised to the exponent.

It is now not difficult to guess what kind of theory of particles and gravitation could resolve the large number paradox: It is necessary that the mass of particles be expressed in terms of the Planck mass in conjunction with the probability of a tunnel transition.¹²⁾

For example, one could imagine an expression of the form

$$m_p = \beta \sqrt{\frac{\hbar c}{G}}\, e^{-4\pi^2},$$

where $\beta$ is a dimensionless factor; numerically, it is found that $\beta = 0.011$; $e^{-4\pi^2} = 7 \cdot 10^{-18}$.

How could one obtain such an expression or an analogous one? Suppose that in Minkowski space there exist several species of massless particles with spin $\frac{1}{2}$. Such particles have a definite helicity (right- or left-handed). Neutrinos are an example of such particles.

In curved (Riemannian) spacetime one can as before distinguish particles with right- and left-handed helicities since the space is topologically similar to Minkowski space.

Suppose however that with a definite probability there can be perturbations in the vacuum that change its topology. To form a mental picture, we can imagine "arms" such that a right-handed particle goes in and a left-handed particle comes out of them.

Then there is a definite probability for transitions between right- and left-handed particles, which effectively corresponds to the presence of rest mass for the particles.

In such a theory, the mass of the particles depends for dimensionless reasons on the quantum fluctuations of spacetime and, therefore, must be of the order of the Planck mass, $10^{-5}$ g $= \sqrt{\hbar c/G}$. But because of the topology change, we must expect the appearance of an exponential small dimensionless parameter. We have written here $e^{-4\pi^2}$ simply as an example of such a parameter; I know of no definite calculations of the formation of such "arms", nor any theory of the quantities $\beta$ for different particles. It may prove harder to understand the existence of the massless neutrinos than to describe the baryon masses.

Let us not get bogged down in making conjectures a about theories that have yet to be developed. Only one comment is important: Dirac's comment was correct as applied to theories of perturbation type, but it does not remain true for theories in which one considers tunnel transitions and topology changes. Thus, we can eliminate one, if not the most serious, hindrance in the path leading to the development of a unified theory of particles and gravitation.

---

¹²⁾ For a different approach to this question, see Markov's paper.[37]

[1] S. W. Hawking, Nature, **248**, 30 (1974).

[2] V. P. Frolov, Usp. Fiz. Nauk **118**, 473 (1976) [Sov. Phys. Usp. **19**, 244 (1976)].

[3] S. W. Hawking, Preprint OAP-420, California Institute of Technology, Pasadena, California (1975).

[4] S. W. Hawking, Phys. Rev. D **14**, 2460 (1975).

[5] V. L. Ginzburg and L. M. Ozernoĭ, Zh. Eksp. Teor. Fiz. **47**, 1030 (1964) [Sov. Phys. JETP **20**, 689 (1965)].

[6] A. G. Doroshkevich, Ya. B. Zel'dovich, and I. D. Novikov, Zh. Eksp. Teor. Fiz. **49**, 170 (1965) [Sov. Phys. JETP **22**, 122 (1966)].

[7] Ya. B. Zel'dovich, Zh. Eksp. Teor. Fiz. **42**, 641 (1962) [Sov. Phys. JETP **15**, 446 (1962)].

[8] L. D. Landau and E. M. Lifshitz, Teoriya Polya (The Classical Theory of Fields, Pergamon Press, Oxoford, 1975), Nauka, Moscow (1973).

[9] L. Parker, Phys. Rev. Lett. **21**, 562 (1968).

[10] Ya. B. Zel'dovich and L. P. Pitaevskii, Comm. Math. Phys. **23**, 185 (1971).

[11] I. D. Novikov, Zh. Eksp. Teor. Fiz. **71**, 393 (1976) [Sov. Phys. JETP **44** 207 (1976)].

[12] A. A. Starobinskiĭ, private communication (1976).
[13] M. A. Markov, Giperony i K-Mesony (Hyperons and K Mesons), Fizmatgiz, Moscow (1958).
[14] Ya. B. Zel'dovich, Preprint [in Russian], IPM Akad. Nauk SSSR, No. 57, Moscow (1976); Pis'ma Zh. Eksp. Teor. Fiz. **24**, 29 (1976) [JETP Lett, **24**, 25 (1976)].
[15] V. L. Ginzburg and V. P. Frolov, Pis'ma Astron. Zh. **2**, 474 (1976) [Sov. Astron. Lett. **2**, (1976)].
[16] R. P. Feynman, Acta Phys. Pol. **24**, 697 (1963).
[17] J. R. Oppenheimer and G. M. Volkoff, Phys. Rev. **55**, 374 (1939).
[18] Ya. B. Zel'dovich and I. D. Novikov, Astron. Zh. **43**, 758 (1966) [Sov. Astron. **10**, 602 (1967)]; Teoriya Tyagoteniya i Evolyutsiya Zvezd (Theory of Gravitation and Evolution of Stars), Nauka, Moscow (1971), p. 458; in: Nestatsionarnye Yavleniya v Galaktikakh (Nonstationary Phenomena in Galaxies, Proceedings of 29th IAU Symposium, 1966), Erevan (1968), p. 280.
[19] E. M. Lifshitz, Zh. Eksp. Teor. Fiz. **16**, 587 (1946).
[20] I. D. Novikov, Zh. Eksp. Teor. Fiz. **46**, 686 (1964) [Sov. Phys. JETP **19**, 467 (1964)].
[21] Ya. B. Zel'dovich and I. D. Novikov, Relyativistskaya Astrofizika (Relativistic Astrophysics, Vol. I: Stars and Relativity, University Chicago Press, Chicago (1971); Vol. II: The Universe and Relativity, Publication of translation by University of Chicago Press announced), Nauka, Moscow (1967).
[22] Ya. B. Zel'dovich, Preprint [in Russian], IPM Akad. Nauk SSSR, No. 98, Moscow (1976); Zh. Eksp. Teor. Fiz. **72**, 18 (1977) [Sov. Phys. JETP **45**, 9 (1977)].
[23] S. W. Hawking, Phys. Rev. D **13**, 191 (1976).
[24] P. I. Fomin, Dokl. Akad. Nauk SSSR Ser. A. 831 (1971).
[25] Ya. B. Zel'dovich and I. D. Novikov, Stroenie i Evolyutsiya Vselennoĭ (Structure and Evolution of the Universe), Izd-vo Akad. Nauk SSSR, Moscow (1975).
[26] A. M. Polyakov, Phys. Lett. B **59**, 82 (1975).
[27] A. A. Belavin *et al.*, Phys. Lett. B **59**, 85 (1975).
[28] G. t'Hooft, Phys. Rev. Lett. **37**, 8 (1976).
[29] J. A. Wheeler, Einstein's Vision, Springer, Berlin (1968) [Russian translation, Mir, Moscow, 1975].
[30] P. A. M. Dirac, Nature **139**, 323 (1937).
[31] E. Schrödinger, Physica **6**, 899 (1939); Proc. R. Irish. Acad. **46**, 25 (1940).
[32] Ya. B. Zel'dovich and A. A. Starobinskiĭ, Zh. Eksp. Teor. Fiz. **61**, 2161 (1971) [Sov. Phys. JETP **34**, 1159 (1972)].
[33] V. N. Lukash and A. A. Starobinskiĭ, Zh. Eksp. Teor. Fiz. **66**, 1515 (1974) [Sov. Phys. JETP **39**, 742 (1974)].
[34] Ya. B. Zel'dovich, I. D. Novikov, and A. A. Starobinskiĭ, Zh. Eksp. Teor. Fiz. **66**, 1897 (1974) [Sov. Phys. JETP **39**, 933 (1974)].
[35] I. W. Roxburgh, Nature, **268**, 504 (1977).
[36] W. Eichendorf and M. Reinhardt, Z. Naturforsch. Teil A **32**, 532 (1977).
[37] M. A. Markov, Zh. Eksp. Teor. Fiz. **51**, 878 (1966) [Sov. Phys. JETP **24**, 584 (1967)].

Translated by Julian B. Barbour