Amplification of gravitational waves in an isotropic universe

L. P. Grishchuk

State Astronomical Institute
(Submitted April 1, 1974)
Zh. Eksp. Teor. Fiz. 67, 825–838 (September 1974)

It is shown that weak gravitational waves in a nonstationary isotropic universe can be amplified to a greater degree than indicated by the adiabatic law. It is a necessary (but not a sufficient) condition for the amplification that there should exist such a stage in the evolution of the universe when the characteristic time for change in the background metric is less than the period of the wave. In an expanding universe the wave is the more amplified the more strongly does the rate of evolution of the universe differ from the one which is dictated by matter with the equation of state $p = \epsilon/3$, and the earlier the wave had been "started." The superadiabatic amplification of gravitational waves denotes the possibility of creation of gravitons. An exceptional position is occupied by the "hot" isotropic universe with $p = \epsilon/3$, in which the superadiabatic amplification of gravitational waves and the production of gravitons is impossible. An estimate is made of the converse reaction of the gravitons on the background metric. Apparently, the production of gravitons forbids at least those of the iostropic singularities near which $p > \epsilon/3$.

An amazing property of the Universe surrounding us is its homogeneity and isotropy on a large scale. One reason for being surprised is that the universe is adequately described by a rather special solution of a complicated system of Einstein's gravitational equations–the Friedmann solution. It is difficult to imagine that the universe, even in its youth, would not "have a desire" to be more anisotropic and inhomogeneous. And yet as we increase our knowledge this epoch is pushed back to evermore earlier stages of its development. It remains to assume that there exist mechanisms for the smoothing out of anisotropy and inhomogeneities, the effectiveness of which increases without limit as a singularity is approached, so that only that solution turns out to be free of contradictions which in the main is homogeneous and isotropic practically from the instant of singularity. No less surprising and requiring an explanation is the fact of the existence of the 3° K residual long wavelength radiation which apparently provides evidence that near the singularity matter was described by the equation of state $p = \epsilon/3$ and not by some other one.

Apparently Misner was the first to begin a systematic search for mechanisms of smoothing out the "random" initial inhomogeneities and anisotropy. He proposed to utilize for this purpose the concept of neutrino viscosity[1] and the so-called phenomenon of mixing[2,3]. Unfortunately, a detailed investigation of these processes has shown that the domain of their applicability is not great and they can not solve the problem that was posed[4-6]. A more universal and promising mechanism appears to be the one proposed by Zel'dovich involving the gravitational effect of particles created in an external gravitational field[7,8]. The production of elementary particles in a gravitational field had also been considered earlier (cf., for example, the bibliography given in the paper by Zel'dovich and Starobinskiĭ[8]), but primarily in connection with an attempt to explain the origin of matter surrounding us. The production of particles in a homogeneous isotropic cosmological model have been studied in particularly great detail by Parker[9-11] who has also investigated some general principles connected with this phenomenon.

The quantum problem of the production of particles in an external gravitational field begins with the solution of the corresponding classical wave equation. If there exists a process for the amplification of the classical wave, then in the quantum theory one should expect the possibility of production of particles. We shall consider weak classical gravitational waves in an isotropic nonstationary universe with a flat space and we shall show that the waves can be amplified over and above the adiabatic law.

Since the metric of an isotropic universe is conformally Euclidean, then an important property of the equations which govern free wave fields is their conformal invariance. The simplest generally covariant generalizations of the equations of the electromagnetic and massless spinor fields are conformally invariant[12]. Arguments are given favoring the choice of an equation for the massless scalar field also in a conformally invariant form[13]. This property of the equations is physically embodied into the impossibility of a superadiabatic amplification of waves and of creation of corresponding massless particles[13-15, 9-11].

For weak gravitational waves there is no lack of uniqueness in the choice of equations. Gravitational wave equations are a direct consequence of the Einstein equations, and in order to alter them one would have to alter the Einstein equations. The equations for weak gravitational waves in a filled universe are not conformally invariant and as a result of this there exists a possibility of amplification of gravitational waves and the production of gravitons. The only exception is the isotropic universe filled with matter subject to the equation $p = \epsilon/3$. In this case the scalar curvature R is equal to zero and the wave equation is reduced by a conformal transformation to the equation in a flat universe, where, evidently, no alteration in the properties of the wave occurs.

A necessary (but not a sufficient) condition for the amplification or weakening of the wave is the existence of such a stage in the evolution of the universe when the wave ceases to be a high frequency one (cf., Secs. 1 and 2). In other words, during that epoch the rate of evolution is so great that the characteristic time for the change in the background metric is less than the period of the wave. In an expanding Friedmann universe this condition is satisfied near the singularity. It turns out that the gravitational wave becomes the more strongly amplified the more the nature of the evolution of the universe differs from the one which is dictated by matter subject to the equation of state $p = \epsilon/3$, and the earlier this wave had been "started" (cf., Sec. 3). In terms of quantum physics this means that gravitons are produced with particularly great intensity near the singularity (more accurately, near the Planck instant of time $t_p = (G\hbar/c^5)^{1/2}$), in universes which do not have a "hot" beginning ($p \neq \epsilon/3$). One can make an estimate of the nature of the reaction of the gravitons being produced on the background metric (cf., Sec. 4). Apparently, the consequences of the production of gravitons are particularly catastrophic for models with an equation of state stiffer than $p = \epsilon/3$ (i.e., for models with the equation of state $p > \epsilon/3$). There are grounds for assuming that in this case production of gravitons ceases only after the law of expansion takes on a form characteristic of a "hot" universe ($p = \epsilon/3$).

Having investigated effects associated with the production of particles in an anisotropic gravitational field Zel'dovich arrived at the conclusion that anisotropic cosmological models end in "suicide" due to the "first viscosity of the vacuum"[7]. Using this terminology it is apparently possible to state that because of the gravitons produced as a result of the "second viscosity of the vacuum" at least some of the isotropic cosmological models (those with $p > \epsilon/3$) end in "suicide."

1. WEAK GRAVITATIONAL WAVES ON AN ARBITRARY BACKGROUND

One arrives at a physically perceived concept of gravitational waves by regarding them as small corrections oscillating in space and time to a certain background metric of space-time. In this case gravitational waves are completely analogous to electromagnetic ones. They are of a transverse nature, propagate with the speed of light, their rays are the light (isotropic) geodesics of the background space-time, the totality of such waves behaves as matter with the equation of state $p = \epsilon/3$, etc.[16-19]

Weak gravitational waves are an inevitable consequence of the Einstein theory of gravitation. Indeed, let the space-time metric $g_{\mu\nu}$ be representable in the form of a sum of a background metric $\gamma_{\mu\nu}$ and small corrections $h_{\mu\nu}$: $g_{\mu\nu} = \gamma_{\mu\nu} + h_{\mu\nu}$. One can introduce auxiliary quantities $\psi_{\mu\nu} = h_{\mu\nu} - 1/2\gamma_{\mu\nu}h$ and such a system of coordinates in which the gauge conditions

$$\psi^{\mu\nu}{}_{;\nu}=0. \qquad (1)$$

are satisfied. Then in the approximation linear in the quantities $\psi_{\mu\nu}$ and their derivatives the Einstein equations in a vacuum are reduced to the system of wave equations

$$\psi_{\mu\nu;\alpha;}{}^{\alpha} - 2R^{(0)}_{\alpha\mu\nu\beta}\psi^{\alpha\beta}=0. \qquad (2)$$

In (1) and (2), the operations of covariant differentiation and transposition of indices, and also the construction of the curvature tensor of index (0), are carried out utilizing the background metric $\gamma_{\mu\nu}$.

The gauge conditions (1) remain unaltered under the small transformations $\bar{x}^\mu = x^\mu + \xi^\mu$ which satisfy the equations

$$\xi_{\mu;\alpha;}{}^{\alpha} + \xi^{\alpha}R^{(0)}_{\alpha\mu}= 0.$$

and this reduces in vacuum to

$$\xi_{\mu;\alpha;}{}^{\alpha}=0. \qquad (3)$$

Utilizing this additional degree of arbitrariness one can reduce the number of unknown functions and simplify the equations. This can be particularly easily achieved in the case of a flat background universe. If $\gamma_{\mu\nu} = \eta_{\mu\nu}$ is the metric of a Minkowski universe, then in addition to conditions (1) one can achieve, for example, that the conditions $\psi = 0, h_{\mu\nu}u^\nu = 0$ are satisfied where u^ν is an arbitrary covariantly constant 4-vector[18]. In particular, if $u^\nu = (1,0,0,0)$, then $h_{0\alpha} = 0$, while the field equations reduce to

$$h_{ik,\alpha,}{}^{\alpha}=0.$$

In the case of a warped vacuum background one can partially utilize the allowable arbitrariness (3) and obtain $\psi = 0$, as a result of which the equations assume the form[18]

$$h_{\mu\nu;\alpha;}{}^{\alpha}-2R^{(0)}_{\alpha\mu\nu\beta}h^{\alpha\beta}=0, \quad h^{\nu}_{\mu;\nu}=0, \quad h=0. \qquad (4)$$

The question of how one can get rid of the remaining nonphysical degrees of freedom has not been solved.

If the universe is filled with matter with the energy-momentum tensor $T_{\mu\nu}$, then the right-hand side of (2) does not vanish. It is natural to define the free (sourceless) gravitational waves by the requirement that the perturbed value of the tensor $T^{\nu(1)}_\mu$ should be equal to zero. This requirement corresponds to the intuitive concept that a free gravitational wave is an alteration in the metric, and not in the matter which fills the universe. In the case $T^{\nu(1)}_\mu = 0$ the Einstein equations reduce to $R^{\nu(1)}_\mu = 0$ and taking (1) into account assume the form

$$h^{\alpha}_{\mu\nu;\alpha;} - 2R^{(0)}_{\alpha\mu\nu\beta}h^{\alpha\beta}=0. \qquad (5)$$

Comparing this with (4) we see that with such a definition of a free gravitational wave the field equations in an empty and a filled form coincide. At least in some special cases one can achieve that the whole set of requirements attainable in a flat background universe are satisfied, and in particular

$$T_\mu{}^{\nu(1)}=0, \quad \psi^{\mu\nu}{}_{;\nu}=0, \quad \psi=0, \quad h_{\mu\nu}u^{\nu}=0. \qquad (6)$$

This holds, for example, in the case of free gravitational waves on an isotropic (Friedmann) background where the role of u^ν is assumed by the (noncovariantly constant) 4-vector of the unperturbed velocity of matter. We note that Misner, Thorne and Wheeler[18] consider it impossible to satisfy globally the conditions $\psi_{\mu\nu}u^\nu = 0$ at the same time as $\psi^{\mu\nu}_{;\nu} = 0$, and explain this by the fact that the background curvature in the general case "does not permit the vector field u^ν to be covariantly constant." In actual fact the existence of a covariantly constant vector field although it is a sufficient condition is not a necessary condition for the possibility of the choice $\psi_{\mu\nu}u^\nu = 0$ together with $\psi^{\mu\nu}_{;\nu} = 0$.

We shall speak of a gravitational wave as being weak if its relative amplitude h^ν_μ is small compared to unity; we shall say that it is short if its length λ is as small compared to the characteristic radii of curvature L of the background space; we shall say that it is high frequency if its period T is small compared to the characteristic time θ for the change in the background metric. Short high frequency waves are propagated with the velocity c, i.e., for them λ and T are related by the equation $\lambda/T = c$. The definitions introduced above have a relative character, they relate the properties of the wave with the properties of the background space and time.

The structure and the evolution of the background universe are determined by the matter of filling the universe and to a greater or lesser degree by the energy and the pressure of the gravitational waves themselves. It is important that in the different regions of background space or in the process of its evolution the corrections to the background metric which we identify with gravitational waves generally speaking cease to be oscillating and small. In other words, the gravitational waves cease to be short, high frequency or even weak. If in two regions of space-time the wave is short and of high frequency, and somewhere in the intermediate region, say, the property of it being a high frequency wave is violated and instead the inequality $T \gg \theta$ is satisfied, then in this region the corrections to the metric lose the usual attributes of a "wave," but we shall as before speak of them as of a wave, albeit a low frequency one.

2. AMPLIFICATION OF GRAVITATIONAL WAVES IN A "DIRECTED" ISOTROPIC UNIVERSE

Weak gravitational waves in an isotropic universe have been investigated in the famous paper of E. Lifshitz in 1946[20]. We restrict ourselves to a background metric with a flat 3-space

$$ds^2=a^2(\eta)\,(d\eta^2-dx^2-dy^2-dz^2). \tag{7}$$

From Eqs. (5), taking conditions (6) into account, we obtain

$$h_i^{k''}+2\frac{a'}{a}h_i^{k'}+a^2g^{lm}h^{k}_{i,l,m}=0, \tag{8}$$

where the prime denotes a derivative with respect to η, while the comma denotes derivatives with respect to the spatial coordinates. According to Lifshitz the wave corrections to the metric can be represented in the form of a sum of term $h_i^k = \nu(\eta)G_i^k$, where G_i^k is the tensor eigenfunction numbered n of the Laplace operator, constructed in accordance with the metric $dl^2 = dx^2 + dy^2 + dz^2$. Then for ν we obtain the equation

$$\nu''+2\frac{a'}{a}\nu'+n^2\nu=0.$$

The number n indicates the spatial periodicity of the wave. Evidently in the case of a flat 3-space any wave corresponding to any $n \neq 0$ is short, since the radius of curvature of the background space is infinite. In the process of the evolution of the background the true wavelength λ (so to speak measured in centimeters) varies proportionally to a, but the ratio between λ and a remains unchanged: $n = 2\pi a/\lambda$. Thus, as a changes the wave does not cease to be short, but the property of the wave of being of high or low frequency can be violated.

The specific form of the function $a(\eta)$ is determined by the matter filling the universe and depends on its equation of state. We consider first the case of the "directed cosmology" (engine-drive cosmology using Wheeler's expression), assuming that a varies in accordance with a previously prescribed law. Let a as $\eta \to -\infty$ and $\eta \to +\infty$ tend respectively to the constant values a_1 and a_2, while in the interval it varies in an arbitrary sufficiently gradual manner. We consider some one component h (we do not show the indices) of a monochromatic wave field depending on η and x. In a manner analogous to the way this was done in[8] we use the modification of the Lagrange method[21]. We seek the solution of Eq. (8) in the form

$$h=\frac{A(\eta)}{a}e^{-in(\eta-x)}+\frac{B(\eta)}{a}e^{in(\eta+x)}$$

with the additional condition $A'e^{-in\eta} + B'e^{in\eta} = 0$.

Instead of one equation of a second order we obtain two equations of first order:

$$A'=\frac{1}{2n}i\frac{a''}{a}(A+Be^{2in\eta}),\quad B'=-\frac{1}{2n}i\frac{a''}{a}(B+Ae^{-2in\eta}). \tag{9}$$

A consequence of these equations is the equation $|A|^2 - |B|^2 = \text{const}$.

As $\eta \to -\infty$ and $\eta \to +\infty$ A and B tend to the constant values A_1, B_1 and A_2, B_2. The characteristic time for a change in the background metric $\theta = a/a'$ as $\eta \to \pm\infty$ is much larger than the period of the wave $T = n$, and we have here short high-frequency waves with an adiabatically varying amplitude $h \sim \text{const}\cdot a^{-1}$ [20]. The constants A_1, B_1 and A_2, B_2 are connected by the equation $|A_1|^2 - |B_1|^2 = |A_2|^2 - |B_2|^2$.

As can be seen from (9), A and B are strictly constant only when $a'' = 0$. If $a'' \neq 0$ and for $\eta \to -\infty$ a wave is prescribed propagating in only one direction (for example, $A_1 \neq 0, B_1 = 0$), then in future its A-amplitude varies and, moreover, a wave propagating in the opposite direction appears, a B-amplitude appears. As a result for $\eta \to +\infty$ we obtain an initial wave which has been amplified as compared to the adiabatic law $|A_2|^2 = |A_1|^2 + |B_2|^2 > |A_1|^2$ and a generated wave propagated in the opposite direction. If as $\eta \to -\infty$ there exists a standing wave $|A_1|^2 = |B_1|^2$, then for $\eta \to +\infty$ it remains a standing wave, $|A_2|^2 = |B_2|^2$, and its amplification or weakening depends on the initial phase. After averaging over the phase one always obtains wave amplification.

In principle superadiabatic amplification of a wave occurs for any law of variation of a (except for the case $a'' = 0$), but a significant superadiabatic amplification (by a factor of severalfold) occurs only under such rapid variations of the background metric that $\theta \lesssim T$. In future we shall have in mind just this kind of amplification.

We illustrate what has been said above on a specific example. We introduce the true time t in accordance with the equation $c\,dt = a\,d\eta$ and rewrite the equation for ν in the form

$$\ddot{\nu}+3\frac{\dot{a}}{a}\dot{\nu}+\frac{c^2n^2}{a^2}\nu=0. \tag{10}$$

Let $a = \text{const} = R_1$ for $t < 0$ (region I); $a = vt + R_1$ for $0 \le t \le \tau$ (region (II); $a = \text{const} = R_3$ for $t > \tau$ (region III). In regions I and III the solution has the form

$$\nu_1=A_1\sin\left(\frac{nc}{R_1}t+\varphi_1\right),\quad \nu_3=A_3\sin\left(\frac{nc}{R_3}t+\varphi_3\right), \tag{11}$$

where A and φ are arbitrary constants. In region II the nature of the solution depends on the relationship between $\theta = t + R_1/v$ (the time during which a varies by a factor of severalfold) and the periods T_1 and T_3. In going over from region I into region II we have $\theta_1 = R_1/v$, and in going over from region III into region II we have $\theta_3 = R_3/v$. In both cases $\theta/T = nc/v$. If $nc/v \gg 1$, then

$$\nu_2=A_2\frac{1}{a}\sin\left[\left(\frac{c^2n^2}{v^2}-1\right)^{1/2}\ln a+\varphi_2\right],$$

and we as before have short high frequency waves with adiabatically varying amplitude. But if $nc/v \ll 1$, then

$$\nu_2=C_1a^\alpha+C_2a^\beta, \tag{12}$$

where

$$\alpha=-1-\left(1-\frac{c^2n^2}{v^2}\right)^{1/2}\approx-2,\quad \beta=-1+\left(1-\frac{c^2n^2}{v^2}\right)^{1/2}\approx-\frac{1}{2}\frac{c^2n^2}{v^2}.$$

We note two modes of the solution—the first and the second term in (12). The characteristic time $T_2 = |\nu/\dot{\nu}|$ for the variation of the wave correction is equal to $|\alpha^{-1}a/\dot{a}| \approx \theta$ for the former and $|\beta^{-1}a/\dot{a}| \gg \theta$ for the latter, and in both cases the high frequency condition $\theta \gg T$ has been violated. The phase velocity of propagation defined as $u = \lambda\dot{\nu}/\nu$ for the former mode is larger by a factor of v/cn compared to c, and for the latter it is smaller than c by the same factor.

In order to find the relation between the amplitudes A_1 and A_3 one must join the solutions (11), (12) at the points t = 0, t = τ requiring the continuity of ν and $\dot{\nu}$. We introduce the notation $x = R_1/R_3$. Let the universe contract from R_3 to R_1. Then A_1 is expressed in terms of A_3 and φ_3 in accordance with the formula

$$\left(\frac{A_1}{A_3}\right)^2=\frac{1}{x^4(\beta-\alpha)^2}\left\{2\left(x^{-2\alpha}+x^{-2\beta}-2\frac{c^2n^2}{v^2}x^2\right)\right.$$
$$\left.+2(x^{-\beta}-x^{-\alpha})\sin\left[2\frac{cn}{v}(1-x)+2\varphi_3+\psi\right]\left[x^{-2\alpha}+x^{-2\beta}+2\left(1-2\frac{c^2n^2}{v^2}x^2\right)\right]^{1/2}\right\}, \quad (13)$$

where

$$\operatorname{tg}\psi=\frac{1}{2}\frac{v}{cn}(\beta-\alpha)\frac{x^{-\alpha}+x^{-\beta}}{x^{-\beta}-x^{-\alpha}}$$

We consider two possibilities for the choice of the arbitrary phase φ_3 in the case of which $\sin[2cn(1-x)/v + 2\varphi_3 + \psi]$ is equal to +1 or −1. These possibilities correspond to the transition between the regions III and I in accordance with the two modes of solution (12). We assume that $x \ll 1$ and take into account the fact that $cn/v \ll 1$. Then in the former case from (13) we obtain $A_1/A_3 = 1/x^2$, while in the latter case $A_1/A_3 = 1$. But if the transition between the regions III and I occurred slowly (adiabatically), then the ratio of the amplitudes would be equal to $A_1/A_3 = 1/x$. Thus, compared with an adiabatic transition we have an amplification of the wave in the former case and a weakening of the wave in the latter case[1)].

Small deviations from the extreme choices of the phase considered above behave differently. They alter the coefficient in the law $A_1/A_3 = 1/x^2$ only to a small extent and very rapidly alter the law $A_1/A_3 = 1$ itself. In order to find out what effect occurs in a "typical" transition from region III into I, we must average (13) over φ_3. As a result we obtain $(\overline{(A_1/A_3)^2})^{1/2} = 1/x^2\sqrt{2}$, i.e., we again have amplification.

In an expansion of the universe from R_1 to R_3 the amplitude A_3 is expressed in terms of the initial amplitude A_1 and the phase φ_1 in accordance with the formula which can be obtained from (13) by the replacement $A_1 \leftrightarrow A_3, R_1 \leftrightarrow R_3, \varphi_3 + n\tau/R_3 \rightarrow \varphi_1$. The two extreme cases of the choice of phase φ_1 correspond to $A_3/A_1 = 1$ and $A_3/A_1 = x^2$, where as before we have $x = R_1/R_3 \ll 1$. Averaging over the initial phase φ_1 leads to the equation $(\overline{(A_3/A_1)^2})^{1/2} = 1/\sqrt{2}$, i.e., it leads again to an amplification of the wave compared to the adiabatic law $A_3/A_1 = x$.

The effect of amplification of a gravitational wave can occur only when the condition $cn/v \ll 1$ is satisfied. This condition can also be rewritten in the form $\lambda \gg c(t + R_1/v)$. In this form this condition means that the wavelength must be large compared to the distance to the horizon of the particle whose world line began at $t = -R_1/v$.

We introduce the energy density of gravitational waves in accordance with the definition $\epsilon_g = K\dot{\nu}^2$, where K is a constant coefficient. In regions I and III ϵ_g averaged over a period is equal, respectively, to

$$\varepsilon_{g1}=\frac{1}{2}KA_1^2\frac{c^2n^2}{R_1^2}, \quad \varepsilon_{g3}=\frac{1}{2}KA_3^2\frac{c^2n^2}{R_3^2}.$$

In the case of a slow variation of the background metric, ϵ_g averaged over several oscillations behaves as $1/a^4$, i.e., as matter with the equation of state $p = \epsilon/3$. An adiabatic invariant in this process is the quantity E/ω, where E is the total energy in a volume element, ω is the frequency of wave; $E = \epsilon_g V \sim 1/a$, $\omega \sim 1/a$, E/ω = const. In the case of rapid variation of the background metric (solution (12)) ϵ_g corresponding to the first mode varies as $1/a^6$, i.e., as matter with $p = \epsilon$, while ϵ_g, corresponding to the second mode, behaves as $1/a^2$, i.e., as matter with $p = -\epsilon/3$.

The ratio of the densities ϵ_{g1} and ϵ_{g3} can be expressed in terms of the ratio of the amplitudes A_1 and A_3:

$$\frac{\varepsilon_{g1}}{\varepsilon_{g3}}=\left(\frac{A_1}{A_3}\right)^2\frac{1}{x^2},$$

so that the amplification of the wave implies an increase in the energy density over the adiabatic law $\epsilon_{g1}/\epsilon_{g3} = 1/x^4$.

One should not think that a wave field of any arbitrary nature is subject to amplification. For example, we consider free electromagnetic waves in an isotropic universe. It can be easily verified that the generally covariant Maxwell equations

$$\frac{\partial F_{\mu\nu}}{\partial x^\alpha}+\frac{\partial F_{\nu\alpha}}{\partial x^\mu}+\frac{\partial F_{\alpha\mu}}{\partial x^\nu}=0, \quad (14)$$

$$F_{\mu\nu;}{}^{\nu}=0, \quad (15)$$

written in the metric (7) reduce to (14) and to

$$\eta^{\nu\alpha}\partial F_{\mu\nu}/\partial x^\alpha=0, \quad (16)$$

where $\eta^{\nu\alpha}$ is the metric tensor for a flat universe:

$$ds^2=d\eta^2-dx^2-dy^2-dz^2. \quad (17)$$

Thus, from (14) and (16) we obtain for $F_{\mu\nu}$ the same solutions as in a flat universe. It can be easily shown that the energy density of the electromagnetic field is equal to $\epsilon_e = T_{00}/g_{00} = T_0^0 = a^{-4}\bar{\epsilon}_e$, where $\bar{\epsilon}_e$ is calculated in terms of $F_{\mu\nu}$ with the aid of the metric (17) and after averaging over a period does not depend on the time at all. Consequently, ϵ_e always varies in accordance with the adiabatic law irrespectively of the nature of the variation of the scale factor. In particular, if the evolution of a begins with the constant value R_1 and ends with the constant value R_3, then no matter how complicated is the behaviour of a in the intermediate region we always have $\epsilon_{e1}/\epsilon_{e3} = (R_3/R_1)^4$. For ϵ_g we obtain the same relation only under the condition of an infinitely slow transition between R_1 and R_3. The mathematical reason for this difference consists of the conformal invariance of the Maxwell equations and the conformal noninvariance of the gravitational wave equations (cf., Sec. 4).

3. AMPLIFICATION OF GRAVITATIONAL WAVES IN A FRIEDMAN UNIVERSE

The example considered above of a "directed" universe was artificial. It explained the possibility of amplification of gravitational waves in an isotropic universe and illustrated the difference in the behavior of electromagnetic and gravitational waves. We are particularly interested in the real Friedmann universe in which the scale factor is determined by matter with a sensible equation of state. Let $p = q\epsilon$, where [2] $0 \le q \le 1$. From the Einstein equations for a homogeneous isotropic Universe (7) we obtain in terms of η-time

$$a = a_0\eta^{2/(3q+1)}, \tag{18}$$

or in terms of t-time

$$a = b_0 t^{2/3(q+1)}.$$

We introduce $\mu = a\nu$ and rewrite (8) in the form

$$\mu'' + \mu(n^2 - a''/a) = 0. \tag{19}$$

Substitution of (18) into (19) reduces it to the equation for Bessel functions with the exact solution

$$\mu = \eta^{1/2}(C_1 J_\alpha(n\eta) + C_2 N_\alpha(n\eta)),$$

where $\alpha = 3(1-q)/2(1+3q)$, while C_1 and C_2 are arbitrary constants. For the sake of simplicity we assume that α is nonintegral and rewrite the solution for ν in the form

$$\nu = (C_1 J_\alpha(n\eta) + C_2 J_{-\alpha}(n\eta))/a_0\eta^\alpha. \tag{20}$$

The applicability of a linear approximation assumes that $|\nu| \ll 1$.

The condition that the wave is a high frequency one $a/\dot{a} \gg \nu/\dot{\nu}$ in this case can be rewritten in the form $n\eta \gg 1$. For $n\eta \gg 1$ (20) reduces to the asymptotic expression

$$\nu \approx a^{-1} B \sin(n\eta + \psi), \tag{21}$$

where

$$B = \left(\frac{2}{\pi n}\right)^{1/2} (C_1^2 + C_2^2 - 2C_1C_2 \sin\pi\beta)^{1/2},$$

$$\operatorname{tg}\psi = \frac{-C_1\sin(\pi\beta/2) + C_2\cos(\pi\beta/2)}{C_1\cos(\pi\beta/2) - C_2\sin(\pi\beta/2)}, \qquad \beta = \alpha - \frac{1}{2}.$$

(The parameter β is convenient because it vanishes when $q = 1/3$.) At this stage we have short high frequency waves of amplitude proportional to $1/a$ and with an energy density $\epsilon_g \sim 1/a^4$.

In terms of t-time solution (21) has the form

$$\nu = \mathrm{const}\cdot\frac{1}{a}\exp\left\{inc\int\frac{dt}{a}\right\}.$$

It is obtained from (10) if the term $3\dot{a}\nu/a$ is small compared to the remaining terms. But if the last term in this equation is small, then from $\ddot{\nu} + 3\dot{a}\nu/a = 0$ we obtain two modes analogous to (12). One of them yields $\dot{\nu}_1 = \mathrm{const}\cdot a^{-3}$, while the second one yields $\nu_2 = \mathrm{const}$. A more detailed investigation in this latter case leads to

$$\nu_2 = \mathrm{const}\cdot\left[1 - \frac{1}{2}\frac{c^2n^2}{b_0^2}\frac{(2+\beta)^2}{2\beta+3}t^{2/(2+\beta)}\right].$$

In the case $n\eta < 1$ the high frequency property of the wave is violated, and this creates conditions for its amplification. The difficulty consists of determining the amplitude of the "wave" in this epoch[3]. In order to resolve it we adopt the following approach. We assume that for $\eta \ge \eta_0$ (where $n\eta_0 \gg 1$) the scale factor has the expression (18), while for $\eta < \eta_0$ it is constant: $a = \mathrm{const} = a_0\eta_0^{2/(3q+1)}$. One can easily find the amplitude of the wave for $\eta < \eta_0$ in such a universe with a "break" in the scale factor, but one should take into account the distortions introduced by the break itself. The quantity which is of interest to us is the amplitude of the wave which exists when $\eta < \eta_0$, after subtracting the distortions introduced by the break. We illustrate this on an example. We consider a universe filled with matter subject to the equation of state $p = \epsilon/3$ ($q = 1/3$). In this case $a'' = 0$ and the solution of equation (19) is given by

$$\mu = B\sin(n\eta + \psi), \tag{22}$$

where B and ψ are arbitrary constants. On the added interval of the scale factor, where $a = \mathrm{const} = a_0\eta_0$, the solution of (19) also takes the form (22):

$$\mu = A\sin(n\eta + \varphi), \tag{23}$$

but the amplitude and the phase are distorted by the break and differ from B and ψ. Indeed, a'' becomes infinite at the point $\eta = \eta_0$ so that $a'' = a_0\eta_0\delta(\eta - \eta_0)$ and μ' undergoes at this point a discontinuity equal to the value of μ at $\eta = \eta_0$. This discontinuity leads to the fact that the amplitude and the phase of solutions (22) and (23) do not coincide. A subtraction of these distortions (in the present case a return to B and ψ) is automatically achieved by a matching of μ and μ' of the solutions (22) and (23) at the point $\eta = \eta_0$. We shall utilize this approach also for other values of q. We emphasize that this procedure is not a construction of a solution in a universe with a break in the scale factor (in such a solution μ' must indeed undergo a discontinuity at $\eta = \eta_0$), but a method of choosing the appropriate solution for $\eta < \eta_0$ and thereby a determination of the amplitude of the wave. This approach would not have been needed if there existed a satisfactory description of the energy of the wave not only in the high frequency region ($n\eta \gg 1$), but also in the low frequency region ($n\eta \ll 1$).

Equation (19) has the form of a Schrödinger equation, in which the role of the potential is played by a''/a:

$$U(\eta) = a''/a = \beta(1+\beta)/\eta^2.$$

The potential falls to zero as $\eta \to \infty$ and reaches a maximum or a minimum value (depending on the sign of a'') at $\eta = \eta_0$. To the left of this point it vanishes identically. The solution of equation (19) for $\eta \ge \eta_0$ has the form (20), while for $\eta < \eta_0$ it has the form (23). A continuous joining of μ and μ' at the boundary $\eta = \eta_0$ enables us to express C_1, C_2 in terms of A and φ, and, in the final analysis, to relate A to the value of B from formula (21). Just as in Sec. 2, B can be expressed in terms of A and φ (for $\beta = 0$ the term depending on the phase disappears), so that in an expanding universe both an amplification of the wave ($B > A$) and a weakening of it ($B < A$) is possible. After averaging over the phase we obtain the following connection between the amplitudes B and A:

$$\overline{\left(\frac{B}{A}\right)^2} = \frac{\pi(n\eta_0)}{4\cos^2\pi\beta}\left\{\left[1 + \left(\frac{\beta}{n\eta_0}\right)^2\right]\left[J^2_{1/2+\beta} + J^2_{-1/2-\beta} + 2\sin\pi\beta J_{1/2+\beta}J_{-1/2-\beta}\right]\right.$$
$$+ 2\frac{\beta}{n\eta_0}\left[J_{1/2-\beta}J_{-1/2-\beta} - J_{1/2+\beta}J_{-1/2+\beta} + \sin\pi\beta\left(J_{1/2+\beta}J_{1/2-\beta} - J_{-1/2-\beta}J_{-1/2+\beta}\right)\right]$$
$$\left. + \left[J^2_{1/2-\beta} + J^2_{-1/2+\beta} - 2\sin\pi\beta J_{1/2-\beta}J_{-1/2+\beta}\right]\right\}, \tag{24}$$

where the Bessel functions are taken for the value of the argument equal to $n\eta_0$. Just as expected, for $\beta = 0$ we obtain $B/A = 1$ for any value of η_0. In this case the wave is neither amplified nor diminished compared to the adiabatic law of variation $\nu \sim 1/a$.

Let us see to what effect we are led by small deviations in the equation of state from $p = \epsilon/3$. Let $|\beta| \ll n\eta_0 \ll 1$. The Bessel functions can be expanded in powers of the argument $n\eta_0$, and then each term of this series can be expanded in powers of the parameter β. First of all we note that if for the time being we leave aside the term

$$\left(\frac{\beta}{n\eta_0}\right)^2[J^2_{1/2+\beta}(n\eta_0)+J^2_{-1/2-\beta}(n\eta_0)+2\sin\pi\beta\, J_{1/2+\beta}(n\eta_0)J_{-1/2-\beta}(n\eta_0)], \quad (25)$$

then the remaining expression in curly brackets in (24) is symmetric with respect to the replacement of β by $-\beta$. Consequently, all the terms with odd powers of β in this remaining expression will mutually cancel. The principal term, which does not contain β at all, is equal to $4/\pi n\eta_0$. The first correction to it is proportional to $(n\eta_0)^{-1}\beta^2 ln^2(n\eta_0)$ and turns out to be smaller than the principal term in the expansion (25), and in particular smaller than the quantity $(2/\pi n\eta_0)(\beta/n\eta_0)^2$. Writing out the principal terms in the expansion of (24) we obtain

$$\overline{\left(\frac{B}{A}\right)^2}=1+\frac{1}{2}\left(\frac{\beta}{n\eta_0}\right)^2.$$

Thus, for small β the function $(B/A)^2$ is symmetric with respect to β and attains at $\beta = 0$ a minimum equal to unity. The coefficient of amplification $K = \overline{(B/A)^2} - 1 = 1/2(\beta/n\eta_0)^2$ is the greater, the more the equation of state differs from $p = \epsilon/3$, the longer is the wave under consideration, and the earlier it had been "started".

4. CONFORMAL INVARIANCE OF WAVE EQUATIONS AND THE PRODUCTION OF GRAVITONS

An important property of homogeneous isotropic metrics is their conformal Euclidean property. This means that the linear element ds can be rewritten in the form $ds = \Omega d\bar{s}$, where $d\bar{s}$ is a linear element of a flat universe, while Ω is a conformal factor which depends, generally speaking, on all the coordinates. In the case of the metric (7) of interest to us the conformal factor coincides with a. The property of being conformally Euclidean is important because the equations for certain free fields written in terms of a homogeneous isotropic metric may be reduced by a conformal transformation of the metric and of the field functions to equations in a flat universe. This applies, for example, to the conformally invariant generalization of the equation for a massless scalar field and to Maxwell equations (14) and (15)[13, 23]. It is evident that in a flat universe with wave solutions nothing happens, and the initial field functions differ only by being multiplied by conformal factor taken to the appropriate power. Consequently, processes of nonadiabatic amplification of waves and production of particles are impossible[9].

Equations for weak gravitational waves on an isotropic background are, generally speaking, not conformally invariant. Indeed, if they were such, then by a transition to the metric (17) and the field function $\sigma = a^r\nu$, where r is a number, we should have obtained from equations (5) instead of (8) an equation of the form $\sigma'' + n^2\sigma = 0$. The best situation that we could obtain by a choice of the number $r = 1$ ($\mu = a\nu$), is equation (19), in which there is an extra term a''/a. Generally speaking, it is not equal to zero, since

$$a''/a=-{}^1\!/_6 a^2R^{(0)}={}^1\!/_3 a^2T^{(0)},$$

where $R^{(0)}$ is the background scalar curvature, while $T^{(0)}$ is the trace of the energy-momentum tensor. Thus, only in the case $T^{(0)} = 0$ and, consequently, for $a = a_0\eta$ + const the equations turn out to be conformally invariant with respect to the transformation with the conformal factor $\Omega = a$.

Parker asserts[9] that massless particles of any arbitrary spin $s > 0$ are not produced in a nonstationary isotropic universe, referring with respect to this to the result due to Penrose[13], according to which the simplest generally covariant generalizations of the field equations for $s > 0$ are all conformally invariant. But Penrose's result for $s = 2$ refers to a field in a vacuum, where $T_{\mu\nu} \equiv 0$ and, therefore, $T \equiv 0$. In a filled isotropic universe, as we can see, the equations for weak gravitational waves are conformally invariant only in the case $T^{(0)} = 0$ and, consequently, only in this case amplification of waves and production of gravitons are not possible. We note that equation (10) coincides exactly with the equation for a massless scalar field in its simplest generally covariant (conformally noninvariant!) generalization. For $T = 0$ Parker's result[9] concerning the absence of production of massless particles is valid.

The correct method of taking into account the reaction of the particles being produced on the metric requires the solution of the self-consistent quantum problem. Certain aspects of this process can be modelled with the aid of the solution of the gravitational equations in a filled universe with an energy-momentum tensor containing viscous terms, where it is assumed that the coefficient of viscosity is switched on at a certain instant of time. We quote, for example, certain results which refer to modelling the production of particles in an anisotropic gravitational field. The production of particles in a vacuum Kasner model[8] and their reaction on the metric can be approximately described with the aid of joining at a certain instant of time t_0 two solutions of Einstein's equations. For $t < t_0$ this is a vacuum Kasner solution, for $t > t_0$ it is a solution with a Kasner metric $ds^2 = c^2dt^2 - a^2dx^2 - b^2dy^2 - c^2dz^2$ and with an energy-momentum tensor with first viscosity. It is assumed that the coefficient of viscosity λ is switched on sufficiently gradually at the instant t_0. In the case of a sufficiently gradual switching on, one can achieve, for example, that at that instant of time the values of the metric and of its first derivatives would be continuous, and the values of the density and the pressure and their first derivatives would be equal to zero. Since in virtue of the Einstein equations we have

$$\dot{\varepsilon}+(\varepsilon+p)\frac{(abc)^{\cdot}}{abc}=\lambda\pi_{ik}\pi^{ik}, \quad (26)$$

where π_{ik} is the tensor of the anisotropy of the deformation, then for $t > t_0$ $\epsilon > 0$ and $p > 0$ will appear[4)]. Due to the gravitational effect of the matter that has appeared the expansion tends to become isotropic and the right-hand part of (26) disappears. As a result one obtains the isotropically expanding Friedmann model filled with matter. Of course, the artificiality of such a construction consists of an arbitrary prescription of the law for the variation of the coefficient of viscosity[5)].

In an analogous manner we can attempt to model the production of gravitons in an isotropic universe with the aid of the second viscosity. Let the energy-momentum tensor consist of two terms: $T_{\mu\nu} = T_{\mu\nu(1)} + T_{\mu\nu(2)}$, where $T_{\mu\nu(1)}$ corresponds to matter with the equation of state $p_1 = q\epsilon_1$, $T_{\mu\nu(2)}$ corresponds to matter with the equilibrium pressure $p_2 = \epsilon_2/3$ and second viscosity ζ. We assume that individually the following equations are satisfied $T^{\nu}_{\mu(1);\nu} = 0$, $T^{\nu}_{\mu(2);\nu} = 0$. Then the Einstein equations in an isotropic universe (7) reduce to the system

$$\varepsilon_1=\frac{C_1}{a^{3(1+q)}}, \quad \varkappa(\varepsilon_1+\varepsilon_2)=\frac{3a'^2}{a^4}, \quad \varepsilon_2 a^4=3\int \zeta a a'^2\, d\eta, \tag{R}$$

where κ is the Einstein gravitational constant, C_1 = const. It is sensible to assume that ζ has the form $\zeta = \chi(\eta)|a''/a|$, where $\chi(\eta)$ increases sufficiently gradually and then falls off in the neighborhood of a certain η_0. Let us examine the limiting case of a δ-like switching on of viscosity. We write $\zeta = \pm\chi\delta(\eta-\eta_0)\, a''/a$, where χ is a number, the sign plus or minus is chosen respectively for $a'' > 0$ or $a'' < 0$. Then $\epsilon_2 = \pm 3\chi a^{-4}(a''a'^2)|_{\eta=\eta_0}$, and for the determination of a we obtain the equation

$$3a'^2=\varkappa[C_1 a^{2\beta/(1+\beta)}\pm 3\chi(a''a'^2)|_{\eta=\eta_0}], \quad \beta=\frac{1-3q}{1+3q}. \tag{S}$$

From this equation it can be seen that for $\beta < 0$ ($p_1 > \epsilon_1/3$) and sufficiently large $\eta > \eta_0$ we have $a \approx k_1\eta$, where $k_1 = [\pm\kappa\chi(a''a'^2)|_{\eta=\eta_0}]^{1/2}$, i.e., the law for the expansion of the universe with $p = \epsilon/3$.

In conclusion we note that if the gravitons were created near the Planck instant of time $t_p \approx 10^{-43}$ sec ($\lambda \sim ct_p \sim 10^{-33}$ cm) and after a period of mutual transformations involving other particles have survived until the present, then they must now exist in the form of gravitational waves in the range of wavelengths of fractions of a centimeter. In principle such waves are amenable to experimental discovery, for example with the aid of the technique of electromagnetic detection[27].

The author is grateful to A. A. Starobinskiĭ and particularly to Ya. B. Zel'dovich for useful discussions and for an attentive reading of the manuscript.

1) Naturally, the wave must remain "weak" also after amplifciation in order that linear theory within the framework of which we are operating should be applicable.

2) In fact, the relations obtained below are also valid for $-1/3 < q < 0$.

3) An analogous difficulty is encountered in the quantum problem of the production of particles [9,14,15]. In the classical problem one must somehow "brake" the evolution of the universe in order to give the possibility for the wave to become "established" and to determine its amplitude, in the quantum problem one has to "brake" the evolution, in order to have time to count the particles the number of which is continuously varying. Quantization of weak gravitational waves in a closed Friedmann model was considered in [22].

4) In this case there exists, of course, an interval of time during which the condition of energy dominance $|T_0^0| \geqslant T_i^h$ [24,25] is not satisfied.

5) Recently the self-consistent problem of the production of particles (not of gravitons) and of its reaction on the metric has been solved under certain additional assumptions by Lukash and Starobinskiĭ [26].

1C. W. Misner, Phys. Rev. Lett. 19, 533 (1967).
2C. W. Misner, Phys. Rev. Lett. 22, 1071 (1969).
3V. A. Belinskiĭ, E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk 102, 463 (1970) [Sov. Phys.-Usp. 13, 745 (1971)].
4A. G. Doroshkevich, Ya. B. Zel'divoch and I. D. Novikov, Astrofizika (Astrophysics) 5, 539 (1969).
5A. G. Doroshkevich, V. N. Lukash and I. D. Novikov, Zh. Eksp. Teor. Fiz. 60, 1201 (1971) [Sov. Phys.-JETP 33, 649 (1971)].
6L. P. Grischchuk, A. G. Doroshkevich and V. N. Lukash, Zh. Eksp. Teor. Fiz. 61, 3 (1971) [Sov. Phys.-JETP 34, 1 (1972)].
7Ya. B. Zel'dovich, ZhETF Pis. Red. 12, 443 (1970) [JETP Lett. 12, 307 (197)].
8Ya. B. Zel'dovich and A. A. Starobinskiĭ, Zh. Eksp. Teor. Fiz. 61, 2161 (1971) [Sov. Phys-JETP 34, 1159 (1972)].
9L. Parker, Phys. Rev. 183, 1057 (1969).
10L. Parker, Phys. Rev. D3, 346 (1971).
11L. Parker, Phys. Rev. Lett. 28, 705 (1972).
12R. Penrose, in "Relativity, Groups and Topology" ed. by C. DeWitt and B. DeWitt, N.Y.-London, 1964 (Russ. transl. in Gravitatsiya i topologiya (Gravitation and Topology) Mir, 1966).
13É. A. Tagirov and N. I. Chernikov, Preprint R2-3277, JINR, 1968.
14K. A. Bronnikov and É. A. Tagirov, Preprint R2-4151, JINR, 1968.
15A. A. Grib and S. G. Mamaev, Yad. Fiz. 10, 1276 (1969) [Sov. J. Nucl. Phys. 10, 722 (1970)].
16A. Éinstein, Sobr. Nauchn. Trudov (Collected Scientific Works) 1, Nauka, 1967, p. 631.
17L. D. Landau and E. M. Lifshitz, Teoriya polya (Field Theory), Nauka, 1967 [Pergamon, 1971].
18C. W. Misner, K. S. Thorne and J. A. Wheeler, Gravitation, San Francisco, 1973.
19R. A. Isaacson, Phys. Rev. 166, 1263 (1968).
20E. M. Lifshitz, Zh. Eksp. Teor. Phys. 16, 587 (1946).
21N. Freman and P. Freman, VKB-priblizhenie (Approximation), Mir, 1967.
22A. A. Lyubushin, Teor. Mat. Fiz. 10, 399 (1972).
23R. Penrose, Structure of Space-Time (Russ. Transl.) Mir, 1972.
24S. W. Hawking, Commun, Math. Phys. 18, 301 (1970).
25Ya. B. Zeldovič and L. P. Pitaevskii, Commun. Math. Phys. 23, 185 (1971).
26V. N. Lukash and A. A. Starobinskiĭ, Zh. Eksp. Teor. Phys. 66, 1515 (1974) [Sov. Phys. JETP 39, 742 (1974)].
27V. B. Braginskiĭ, L. P. Grishchuk, A. G. Doroshkevich, Ya. B. Zel'dovich, I. D. Novikov and M. V. Sazhin, Zh. Eksp. Teor. Phys. 65, 1729 (1973) [Sov. Phys. JETP 38, 865 (1974)].

Translated by G. Volkoff.
95