Spectrum of adiabatic perturbations in the universe when there are singularities in the inflaton potential

A. A. Starobinskiĭ

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334, Moscow

(Submitted 9 April 1992)

Pis'ma Zh. Eksp. Teor. Fiz. **55**, No. 9, 477–482 (10 May 1992)

If the potential of the effective scalar which controls the de Sitter (inflationary) stage in the early universe has a singularity consisting of a rounded change in slope, a step of a universal form arises in the spectrum of adiabatic perturbations. Along with this step, there are superimposed modulations. If the singularity in the potential is instead a rounded jump, a hump appears in the spectrum.

One cosmological scenario has an initial stage of a quasiexponential expansion (a de Sitter or inflationary stage). The decay of this stage results in the appearance of a hot Friedmann universe. One of the primary advantages of this scenario, as has been stressed repeatedly, starting in Ref. 1, is that any specific version of it can be refuted or confirmed with the help of observational data on the contemporary universe. The predictions which are easiest to test are those regarding the spectrum and statistics of adiabatic perturbations of the Friedmann metric, which lead to the formation of galaxies and other compact objects in the universe. In all simple versions of the inflationary model it is assumed that the inflationary stage is produced by a single effective scalar field (an inflaton) with a potential $V(\varphi)$. In the inflationary stage, this field is in a slow roll, with

$$|\ddot{\varphi}| \ll 3H|\dot{\varphi}|; \qquad \dot{\varphi}^2 \ll 2V(\varphi), \tag{1}$$

where $H = \dot{a}/a$, $a(t)$ is a scale factor of the isotropic cosmological model, and the dot means differentiation with respect to t. The Fourier components of the gravitational potential $\Phi_k = (2\pi)^{-3/2} \int \Phi(\vec{r})\exp(-i\vec{k}\vec{r})d^3r$ are then δ-correlated Gaussian random quantities ($\langle \Phi_{\vec{k}} \rangle = 0$; $\langle \Phi_{\vec{k}} \Phi_{\vec{k}'} \rangle = \Phi^2(k)\delta^{(3)}(\vec{k} - \vec{k}')$, $k = |\vec{k}|$) with an approximately flat spectrum:

$$\left| \frac{d\ln(k^3\Phi^2(k))}{d\ln k} \right| \ll 1. \tag{2}$$

This problem was first solved completely in Ref. 2 (the results found in the second and third citations of Ref. 2 are identical when expressed in the same way; the result found in the first paper within a numerical coefficient). We might also cite some earlier approaches to this problem.[3]

Under inequalities (1), and with a small spatial curvature—this case sets in right after the beginning of the inflationary stage—we have $H^2 \simeq 8\pi GV/3$ and $\dot{\varphi} \simeq -V'/3H$ ($c = \hbar = 1$, and the prime means the derivative with respect to φ). Conditions (1) can then be rewritten in terms of limitations on the derivatives of V:

$$|V''| \ll 24\pi GV; \quad V'^2 \ll 48\pi GV^2. \tag{3}$$

On the other hand, the absolute value $|V'|$ cannot be too small for the range of φ in which we are interested here in the inflationary stage. This range corresponds to length scales from 1 kpc to 10^4 Mpc today. From the condition that Φ be sufficiently small in this region ($A \leqslant 10^{-3}$ in the notation of Ref. 4) we find

$$|V'| \geq 10^5 (GV)^{3/2}. \tag{4}$$

The model which currently enjoys the greatest popularity is one of cold particles with a total density of matter equal to the critical density (according to the prediction of the inflationary model). For this cold-particle model, prediction (2) agrees well with the shape of the observed correlation function of galaxies (when a nonlinear evolution is taken into account for redshifts $Z \leqslant 10$) for length scales $L = (1\text{–}20) h_{50}^{-1}$ Mpc, where $h_{50} = H_1/50$, and H_1 is the Hubble constant, in units of kilometers per second per megaparsec. At $L = (50\text{–}100) h_{50}^{-1}$Mpc, however, the value of $k^3\Phi^2(k)$ increases, apparently by a factor of at least 3 in comparison with its value in the region $L = (1\text{–}20) h_{50}^{-1}$ Mpc. This effect follows from recent results on the spatial distribution of galaxies ($\propto \Delta\Phi$) at these scales.[5] It also follows from data on the large-scale peculiar velocities of galaxies[6] ($\propto \nabla\Phi$). On the other hand, the distribution of peculiar velocities can be described very accurately as Gaussian.[7] These data refer to the same interval of length scales as that in which the amplitude Φ is greater than that in the cold model with a flat spectrum. The best upper limits on the fluctuations in the temperature of the background radiation, $\Delta T/T$, at the corresponding angular scales (0.25–1°; Ref. 8) would allow a further increase in the amplitude by a factor of about 2. Upper limits on the nondipole anisotropy $\Delta T/T$ at large angles[9] rule out an explanation of this rise based on a scale-invariant spectrum $k^3\Phi^2(k) \propto k^{n-1}$, $n < 1$, which arises in (for example) the "new" inflationary model with a potential $V(\varphi) = V_0 - M^2\varphi^2/2$, $M^2 \sim H_0^2 = 8\pi G V_0/3$, or in a so-called power-law inflation $[a(t) \propto t^p, p > 1]$.

There are several ways to obtain a scale-invariant spectrum of adiabatic perturbations, while retaining their δ-function correlation and Gaussian nature: (A) Abandon the conditions of a slow roll, i.e., (1); (B) introduce several effective scalar fields which can create a de Sitter stage (in taking this approach, we run into the models of double[10] or multiple[11] inflation); (C) keep the initial flat spectrum, (2), but switch from the pure cold-particle model to mixed models. These mixed models would consist of cold particles plus a light neutrino or cold particles plus a cosmological constant, with the resultant density of all types of matter being equal to the critical value. [Where it appears in observables, the initial spectrum $\Phi^2(k)$ is multiplied by the transfer function $c^2(k)$, which arises in the transition from the radiation-dominated stage to the matter-dominated stage at $Z \sim 10^4$, and which depends on the contemporary composition of matter in the universe.]

In the present paper we consider the first of these possibilities. The first differs from the second in that there are no problems in choosing initial values for the scalar field. It differs from the third in that fewer restrictions are imposed on the perturbation spectrum. In order to maintain the agreement with the shape of the correlation function of the galaxies in the region $(1\text{–}20)h_{50}^{-1}$ Mpc, and in order to avoid going beyond the limitations which follow from $\Delta T/T$ at large angles, it is reasonable to assume (in contrast with Ref. 12, where a polynomial potential was discussed) that inequalities (1) and (3) hold everywhere except in a narrow region $\varphi \simeq \varphi_0$ (actually, we are violating only the first inequality). The spectrum then becomes flat far from the point $k_0 = a(t_0)H_0$, where we have $H_0^2 = 8\pi G V(\varphi_0)/3$, where t_0 is the time at which the equality $\varphi = \varphi_0$ holds (for simplicity, we shift the origin of the time scale to satisfy $t_0 = 0$). We set $x = \varphi - \varphi_0$. If $V(x)$ has a singularity in the form of a rounded jump in the second derivative or a weaker singularity near the point $x = 0$, then it follows from the satisfaction of conditions (1) before and after passage through the point $x = 0$ that the perturbation spectrum remains flat.

The minimal local singularity in $V(x)$ sufficient to give rise to a nonflat spectrum is a rounded slope change:

$$\begin{aligned} V(x) = V_0 + v(x), \quad v(x) &\approx A_+ x, \qquad\qquad x \gg x_0, \\ &\approx A_- x, \quad x < 0, \quad |x| \gg x_0, \\ v(0) = 0, \quad A_+ > 0, \quad & A_- > 0, \quad A_+ \neq A_-, \end{aligned} \tag{5}$$

where x_0 is a scale width of the transition region. From the satisfaction of conditions (3) and (4) before and after the transition through the change in slope of the potential, we conclude the following: $\max(A_+, A_-) \ll G^{-1/2} H_0^2$, $\min(A_+, A_-) \gg H_0^3$. If the slow roll is to be disrupted during the transition, we must have $|A_+ - A_-| \gtrsim H_0^2 x_0$. It follows that we have $x_0 \ll G^{-1/2}$ and $x_0 \ll G^{-1/2}$ and $\max(A_+, A_-) \cdot x_0 \ll V_0$. In calculating $a(t)$, we can thus ignore the contribution of v to the potential, and we have $a(t) = a(0)\exp(H_0 t)$ near the transition.

We next consider the case $\min(A_+, |A_+ - A_-|) \gg H_0^2 x_0$, in which the deviation from a flat spectrum is at its greatest. In this limit the spectrum does not depend on the shape of $v(x)$ or the value of x_0; it assumes a universal form. In this case, the region $\Delta x \sim x_0$ is crossed over a time $t_0 \sim x_0 H_0 / A_+ \ll H_0^{-1}$. We thus have

$$\dot{x} = \begin{cases} -A_+/3H_0, & t < 0, \\ -(A_- + (A_+ - A_-)e^{-3H_0 t})/3H_0, & t \geq 0, \end{cases} \tag{6}$$

regardless of the form of $v(x)$.

The equations for small inhomogeneous perturbations of the field φ and for scalar perturbations of the metric can be reduced to a single equation for the gravitational potential[13] Φ or the quantity[14] ξ [in the synchronous frame of reference $\xi = \delta\varphi - \dot{\varphi}/6H(\lambda + \mu)$, where λ and μ are Lifshitz notation for scalar perturbations of the metric[15]]. It is more convenient to use the second of these quantities here. The equation for its Fourier component is

$$\begin{aligned} \ddot{\xi}_k + 3H\dot{\xi}_k + \left(\frac{k^2}{a^2} + m_{eff}^2\right)\xi_k = 0, \\ m_{eff}^2 = \frac{d^2 V}{d\varphi^2} + 8\pi G \frac{\dot{\varphi}}{H}\frac{dV}{d\varphi} + H\left(\frac{\dot{H}}{H^2}\right)^{\cdot} . \end{aligned} \tag{7}$$

Since we have $\dot{H} = -4\pi G\dot{\varphi}^2$, we can replace the quantity $m_{\rm eff}^2$ in (7) by $[3H_0(A_+ - A_-)/A_+]\delta(t)$ in a leading approximation under the condition $k^2 \ll k_0^2(|A_+ - A_-|/H_0^2 x_0)$, by virtue of the inequalities written above (the term $d^2V/d\varphi^2$ makes the basic contribution). The correctly normalized solution for ξ_k at $t<0$, which corresponds to an "in" vacuum as $t \to -\infty$, is

$$\xi_k = \frac{H_0}{\sqrt{2k}} e^{-ik\eta}\left(-\eta + \frac{i}{k}\right), \quad \eta = \int \frac{dt}{a(t)} = -(H_0 a)^{-1}. \tag{8}$$

At $t>0$ we then have

$$\xi_k = \frac{H_0}{\sqrt{2k}}\left(\alpha(k)e^{-ik\eta}\left(-\eta + \frac{i}{k}\right) + \beta(k)e^{ik\eta}\left(-\eta - \frac{i}{k}\right)\right),$$

$$\alpha = 1 + \frac{3i}{2}\left(\frac{A_-}{A_+} - 1\right)\frac{k_0}{k}\left(1 + \frac{k_0^2}{k^2}\right),$$

$$\beta = -\frac{3i}{2}\left(\frac{A_-}{A_+} - 1\right)\exp\left(2i\frac{k}{k_0}\right)\frac{k_0}{k}\left(1 + \frac{ik_0}{k}\right)^2,$$

$$|\alpha|^2 - |\beta|^2 = 1. \tag{9}$$

The quantity $|\beta(k)|^2$ can be interpreted as the number of pairs of scalar particles with momenta $\vec{k}$ and $-\vec{k}$ which are created because of the rapid variation in φ. However, we are interested in only that part of the effect which contributes to a growing adiabatic model of the perturbations. This part is determined by the asymptotic behavior of ξ_k as $t \to \infty$ $(\eta \to 0)$: $\xi_k(\infty) = iH_0/\sqrt{2k^3}(\alpha - \beta)$. Transforming from ξ to Φ in the standard way, using the quantity $h(\vec{r}) = \mu/3$, and assuming $a(t) \propto t^{2/3}$ at present, we find the contemporary perturbation spectrum in the linear stage:

$$\Phi = -\frac{1}{2}h\left(1 - \frac{H}{a}\int_0^t a\,dt\right) = -\frac{3}{10}h,$$

$$k^3 h^2(k) = \frac{18H_0^6}{A_-^2}D^2\left(\frac{k}{k_0}\right)c^2(k),$$

$$D^2 = |\alpha - \beta|^2 = 1 - 3\left(\frac{A_-}{A_+} - 1\right)\frac{1}{y}\left(\left(1 - \frac{1}{y^2}\right)\sin 2y + \frac{2}{y}\cos 2y\right)$$
$$+\frac{9}{2}\left(\frac{A_-}{A_+} - 1\right)^2\frac{1}{y^2}\left(1 + \frac{1}{y^2}\right)\left(1 + \frac{1}{y^2} + \left(1 - \frac{1}{y^2}\right)\cos 2y - \frac{2}{y}\sin 2y\right),$$

$$y = \frac{k}{k_0}, \qquad D(0) = A_-/A_+, \qquad D(\infty) = 1. \tag{10}$$

The function $D(y)$ is a step (with superimposed modulations) with an increase toward large scale values if $A_- > A_+$ and toward small scale values in the opposite case. The shape of this function depends on only the ratio A_-/A_+. This function has two other interesting properties. (1) In the case $A_- > A_+$, the function $D(y)$ falls off slowly toward large values of y (this behavior corresponds to the formation of a large-scale structure in the universe):

$$D \approx 3(A_-/A_+)y^{-1}|\cos y| \text{ for } A_- \gg A_+ \text{ and } 1 \ll y \ll A_-/A_+.$$

(2) In the case $A_- \ll A_+$, the function $D(y)$ falls off $\propto y^2$ toward small values of y as long as the condition $D \gg D(0)$ holds. It has a deep minimum at $y = \sqrt{5A_-/2A_+}$. The value of this function at the minimum is $\sim D(0)(A_-/A_+)^{1/2}$ (a similar effect has been seen in some numerical calculations for polynomial potentials[12]).

With $A_-/A_+ = 3$–3.5 and $k_0^{-1} = (100$–$150)h_{50}^{-1}$ Mpc, the spectrum in (10) is quite successful in explaining the observational data on $\delta\rho/\rho$ and on the peculiar velocities at large scales. It does not contradict the upper limits on $\Delta T/T$ anywhere in the entire angular range[16] (the shift parameter at length scales $< 10h_{50}^{-1}$ Mpc must be $b = 2.2$–2.5). A critical test for this spectrum is the value of $\Delta T/T$ at large angles. In particular, the expected quadrupole anisotropy is $(\Delta T/T)_Q \approx (6$–$7) \times 10^{-6} h_{50}^{-1}$. Since we have $h_{50} \leqslant 1.3$ for the model of cold particles (because of the age of the universe), spectrum (10) can be regarded as definitely refuted if observations reveal $(\Delta T/T)_Q < 3 \times 10^{-6}$. If we wish to explain the observations while adhering to the hypothesis of local singularities in $V(\varphi)$, we must thus switch to the type of singularity which comes in terms of degree of nonanalyticity: a rounded jump in $V(\varphi)$.

If we assume that the size of this jump satisfies $\Delta \ll V_0$ [again, only the first inequalities in (1) and (3) will be violated], and if we assume that the values of $dV/d\varphi$ are equal far from the singularity ($\varphi = \varphi_0$) (the latter assumption is made for simplicity and could easily be abandoned), we find

$$\begin{aligned} V(x) = V_0 + Ax + v(x), \quad & v(x) \approx \Delta/2, \quad x \gg x_0, \\ & \approx -\Delta/2, \quad x < 0, \quad |x| \gg x_0, \\ x = \varphi - \varphi_0, \quad v(0) = 0, & \quad A > 0, \quad \Delta > 0. \end{aligned} \tag{11}$$

It follows from (3) and (4) that we have $H_0^3 \ll A \ll G^{-1/2} H_0^2$. Condition (3) is violated at $|x| \lesssim x_0$ if $\Delta \gtrsim H_0^2 x_0^2$. Furthermore, if the change in $\dot{x}$ is to be important, we must have $\Delta \gtrsim A^2 H_0^{-2}$. It follows that we have $x_0 \ll G^{-1/2}$ and $\Delta \gg H_0^4$. We assume below that the strong inequalities $\Delta \gg H_0^2 x_0^2$ and $\Delta \gg A^2 H_0^{-2}$ hold.

In this case the quantity $m_{\rm eff}^2$ in (7) has a singularity stronger than a δ-function in the formal limit $x_0 \to 0$ (as before, the term $d^2V/d\varphi^2$ makes the basic contribution). The result for the spectrum is thus nonuniversal; i.e., it depends on the shape of $v(x)$ and on the value of x_0. Skipping over the calculations, we write the result. If $A \gg H_0^2 x_0$, then the function $D(k/k_0)$ in (10) (in which we should set $A_+ = A_- = A$) is a plateau with a flat top and superimposed modulations: $D(y) = 3\sqrt{2\Delta}H_0/A\,|\sin y| \gg 1$ at $1 \ll y \ll A/H_0^2 x_0$. At $y \ll 1$, we have $D(y) \propto y^2$ as long as the condition $D \gg 1$ holds. The quantity $D(y)$ has a deep minimum at $y^2 = 5A/H_0\sqrt{2\Delta} \ll 1$. The particular way in which $D(y)$ decreases at $y > A/H_0^2 x_0$ depends on the asymptotic behavior of the quantity $\tilde{v} = \Delta/2 - v(x)$ as $x \to \infty$. If $\tilde{v} \propto \exp(-x/x_0)$ as $x \to \infty$, then

$$D(y) = \frac{3\sqrt{2\Delta}H_0}{A}\frac{1}{\sqrt{1+\delta^2}}|\sin y|, \qquad \delta = \frac{6H_0^2 x_0 y}{A}, \tag{12}$$

at $y \gg 1$ and $D \gg 1$, so we have $D(y) \propto y^{-1}$ at $y \gg A/H_0^2 x_0$. If, on the other hand, we have $\tilde{v} \propto x^{-\gamma}$, then we have $D(y) \propto y^{-\gamma/(\gamma+2)}$.

For $A < H_0^2 x_0$, the plateau at the crest of the hump disappears, we have a maximum value of $D(y) \sim \sqrt{\Delta}/H_0 x_0 \gg 1$ at $y \sim 3$, and the decays on both sides are of the same nature as in the preceding case. If a jump in the potential is stretched out, and the time taken to cross the jump region is longer than H_0^{-1}, the result is precisely the opposite: a well instead of a hump is formed in the spectrum.[17]

[1]A. A. Starobinskiĭ, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 719 (1979) [JETP Lett. **30**, 682 (1979)].
[2]S. W. Hawking, Phys. Lett. B **115**, 295 (1982); A. A. Starobinsky, Phys. Lett. B **117**, 175 (1982); A. H. Guth and S. Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982).
[3]V. N. Dukash, Zh. Eksp. Teor. Fiz. **79**, 1601 (1980) [Sov. Phys. JETP **52**, 807 (1980)]; V. F. Mukhanov and G. V. Chibisov, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 549 (1981) [JETP Lett. **33**, 532 (1981)].
[4]A. A. Starobinskiĭ, Pis'ma Astron. Zh. **9**, 579 (1983) [Sov. Astron. Lett. **9**, 302 (1983)].
[5]S. J. Maddox, G. Efstathiou, W. J. Sutherlan, and J. Loveday, Mon. Not. R. Astron. Soc. **242**, 43 (1990); G. Efstathiou *et al.*, Mon. Not. R. Aston. Soc. **247**, 10 (1990); W. Saunders *et al.*, Nature **349**, 32 (1991).
[6]J. P. Ostriker and Y. Suto, Astrophys. J. **348**, 378 (1990); E. Bertschinger, A. Dekel, S. M. Faber *et al.*, Astrophys. J. **364**, 370 (1990).
[7]L. Kofman, E. Bertschinger, J. Gelb *et al.*, Astrophys. J. (1992) (in press).
[8]P. Meinhold and P. Lubin, Astrophys. J. **370**, L11 (1991).
[9]A. A. Klypin, M. V. Sazhin, I. A. Strukov, and D. P. Skulachev, Pis'ma Astron. Zh. **13**, 259 (1987) [Sov. Astron. Lett. **13**, 104 (1987)]; G. F. Smoot *et al.*, Astrophys. J. **371**, L1 (1991); S. S. Meyer, E. S. Cheng, and L. A. Page, Astrophys. J. **371**, L5 (1991).
[10]L. A. Kofman, A. D. Linde, and A. A. Starobinsky, Phys. Lett. B **157**, 361 (1985); L. A. Kofman and A. D. Linde, Nucl. Phys. B **282**, 555 (1987); J. Silk and M. S. Turner, Phys. Rev. D **35**, 419 (1987).
[11]A. A. Starobinskiĭ, Pis'ma Zh. Eksp. Teor. Fiz. **42**, 124 (1985) [JETP Lett. **42**, 152 (1985)].
[12]H. M. Hodges, G. R. Blumenthal, L. A. Kofman, and J. R. Primack, Nucl. Phys. B **335**, 197 (1990).
[13]M. Sasaki, Progr. Theor. Phys. **70**, 394 (1983).
[14]V. F. Mukhanov, Zh. Eksp. Teor. Fiz. **94**(7), 1 (1988) [Sov. Phys. JETP **67**(7), 1297 (1988)].
[15]E. M. Lifshits, Zh. Eksp. Teor. Fiz. **16**, 587 (1946); L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon, New York, 1976.
[16]N. I. Arzamasova and A. A. Starobinskiĭ, Pis'ma Astron. Zh. (1992) (in press).
[17]D. S. Salopek, J. R. Bond, and J. M. Bardeen, Phys. Rev. D **40**, 1753 (1989); V. F. Mukhanov and M. I. Zelnikov, Phys. Lett. B **263**, 169 (1991).

Translated by D. Parsons