

THE SPECTRUM OF DENSITY PERTURBATIONS IN AN EXPANDING UNIVERSE

JOSEPH SILK*

Astronomy Dept., University of California, Berkeley, Calif., U.S.A.

1. Introduction

Perhaps the most challenging problem confronting a cosmologist is to reconcile the observed large-scale structure of the Universe with the Friedmann-Lemaître cosmological models that have gained such widespread acceptance in recent years (cf. however the alternative viewpoint, as exemplified in this Symposium by Arp and others). In this review, I shall look anew at the spectrum of density inhomogeneities that survive decoupling of matter and radiation at $z \sim 1000$ and provide the primordial fluctuations that can eventually generate galaxies. A closely related matter, that of the associated fluctuations in the background radiation, is discussed elsewhere in this volume by Doroshkevich, Sunyaev and Zel'dovich.

It is apparent from the observed ages, mean densities, and spatial separations of galaxies that we are inevitably confronted with the problem of forming galaxies in the expanding universe, and moreover, that galaxies could not have existed at epochs earlier than those corresponding to a redshift $z \sim 100$ – 1000 . For the purpose of the following discussion, I shall define the initial epoch as $t_i \sim 1$ yr or $z_i \sim 10^8$, corresponding roughly to when the mass of a galaxy is first contained within the particle horizon. Since no instability has been found that can produce large fluctuations from an initially uniform state (Lifshitz, 1946), one must necessarily consider finite amplitude initial conditions.

The various possibilities that arise can be broadly classified into three categories of initial conditions. It is convenient to introduce the parameter $\delta\rho/\rho$, the relative fluctuation in matter density, as a measure of the degree of inhomogeneity on any specified scale. Note incidentally that one must have $|\delta\rho/\rho| \lesssim 1$ at $z \gtrsim 1000$, otherwise premature formation would occur of gravitationally bound systems that can bear little relation to the observed galaxies. Presumably such objects, if present at $z \gtrsim 1000$, would collapse to dense nuclei, and ultimately perhaps form black holes.

1.1. EMPIRICAL INITIAL CONDITIONS

The most direct approach is to postulate an initial amplitude for the density fluctuations that is just adequate to have formed galaxies by the present epoch. One can generalize these conditions somewhat by assuming such fluctuations to be present on all scales. The initial vorticity is taken to be zero.

* Alfred P. Sloan research fellow.

1.2. TURBULENT INITIAL CONDITIONS

Here, one postulates an initial velocity field with non-zero vorticity on all scales, such that $\langle \nabla \cdot \mathbf{u} \rangle_i = 0$. The associated density fluctuations are initially zero, although turbulence is generated on sufficiently small scales ($\lesssim v_0 t$) where dynamical interaction can occur between different eddies. The density fluctuations associated with the turbulence remain small provided that the turbulence is subsonic, or $v_0 \ll c$. This approach has been explored in detail in recent years (e.g., see the papers in this volume by Nariai and Ozernoi).

1.3. CHAOTIC INITIAL CONDITIONS

Finally, one can hypothesize initial conditions of the form $|\delta\varrho/\varrho| \sim 1$ and $\langle v_i^2 \rangle^{1/2} \sim c$. The complexity of this situation lies in the fact it is inherently non-linear, and consequently this approach has not hitherto been developed in any detail (cf. however Misner, 1967; Rees, 1972).

The aim with any set of initial conditions is always to infer the residual spectrum of primordial fluctuations that emerges after decoupling and that will eventually generate galaxies. In the bulk of the ensuing discussion, I shall discuss the consequences of the assumption of empirical initial conditions. This is a phenomenological hypothesis, which appears to provide the simplest and most direct approach to the problem of galaxy formation.

The present discussion is organized as follows. In the following section, the basic dynamical equations are established that govern the evolution of perturbations in a Friedmann-Lemaître universe. I shall work throughout with equations that are essentially post-Newtonian, in order to concentrate in this review on deriving certain physical results. This formalism is appropriate for Friedmann-Lemaître models of arbitrary spatial curvature, provided that one restricts the validity of the discussion to length-scales short compared to the particle horizon.

In Section 3, general solutions are given that describe the evolution of adiabatic perturbations in the matter density, and in Section 4, I study the choice of appropriate initial conditions. The various perturbation modes are compared in Section 5, and Section 6 is devoted to a study of the effect of decoupling on the perturbation spectrum. The scheme that I use for following the evolution of density perturbations through decoupling is based on an extension of the Eddington approximation to the radiative transfer equation, and is strictly valid in both the optically thick and thin limits. A final section summarizes the preceding results, and describes the emergent spectrum at decoupling. Various physical effects that can affect the form of the final spectrum are examined, including the effects of shock formation and Thomson drag.

2. Dynamical Equations

I shall consider a two-fluid system containing matter and radiation, in which the principal coupling is via Thomson scattering by free electrons. The unperturbed fluid

is taken to be the spatially flat Einstein-de Sitter universe, which satisfies the Friedmann equations

$$\varrho a^3 = \text{const}; \quad \frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3} (\varrho + \varrho_r). \quad (1)$$

p is the pressure of both matter and radiation, ϱ the matter density and ϱ_r the radiation mass density. The scale factor a accordingly varies as $t^{2/3}$ if $\varrho \gg \varrho_r$, and as $t^{1/2}$ if $\varrho \ll \varrho_r$ (since $\varrho_r \propto a^{-4}$).

In the comoving coordinate system the perturbed velocity v is of first-order, as also are the perturbed density of matter ϱ_1 ($\equiv \delta\varrho$) and pressure p_1 . The linearized field equations that describe small perturbations of an arbitrary metric tensor h_{ij} relative to a Friedmann universe reduce in lowest order to

$$h_{0,i}^{0,i} = -\frac{8\pi G}{c^2} \delta T_{00}^0, \quad (2)$$

where T_{ij}^j is the perturbed energy-momentum tensor*. In obtaining this result, a generalized de Donder gauge condition (Lanczos, 1925) has been used, and a slow-motion and weak field approximation utilized (cf. Irvine, 1965; Silk, 1966; Layzer, 1968). The perturbed equation of baryon conservation yields

$$\frac{\partial v^\lambda}{\partial x^\lambda} = -\frac{\partial s}{\partial t}, \quad (3)$$

where we have introduced the quantity

$$s \equiv \varrho_1 / \varrho.$$

One now utilizes the conservation equations

$$T_{;i}^{i\lambda} = 0 \quad (4)$$

in order to derive an equation for s .

It is necessary, however, to discuss first the present method for incorporating the radiation field into Equations (2) and (4). One of the principal aims in this discussion will be to follow perturbations through the decoupling era. Hence it is necessary to develop a formalism for incorporating the radiation explicitly into the perturbation equation that remains valid for arbitrary optical depths. Earlier treatment of this problem have been given by Peebles and Yu (1970) and by Bardeen (1968). Peebles and Yu chose to numerically integrate the Boltzmann equation for the photon distribution, whereas Bardeen took some 20 moments of the radiative transfer equation. The complexity of these methods, and in particular their lack of susceptibility to a simple analytical approach, has obscured much of the relevant physics.

Accordingly, I shall develop a somewhat more direct approach to the problem.

* Repeated Latin indices i, j, k etc. are summed over the four coordinates x_1, x_2, x_3 and t ; repeated Greek indices are summed over three spatial coordinates.

The radiative transfer equation can in fact be solved by use of a modified version of the Eddington approximation, an approximation that is valid in both optically thick and thin limits for a radiation field that does not deviate greatly from isotropy (Unno and Spiegel, 1966; Anderson and Spiegel, 1972). At intermediate optical depths, this approximation should be adequate for the present purpose.

To develop an equation for the perturbed mean intensity, one proceeds as follows. The specific radiation intensity $I(\mu_i, x_i, t)$ satisfies the radiative transfer equation

$$\frac{1}{c} \frac{\partial I}{\partial t} + \frac{4}{c} \frac{\dot{a}}{a} I + \mu_i \frac{\partial I}{\partial x_i} + KI = j. \tag{5}$$

This equation is valid in the comoving frame for wavelengths short compared to the particle horizon and for a gray opacity coefficient K (in cm^{-1}); j is the total emissivity (including scattered radiation) and μ_i the direction cosine vector for the angle between x_i and the direction of the radiation beam. The radiation field is anisotropic because of the perturbed velocity field, and to lowest order in v/c , the transformation properties lead to the relations (Thomas, 1930)

$$j' = j(1 - 3\mu_i v_i/c); \quad K' = K(1 + \mu_i v_i/c); \quad J' = J - 2v_i H_i/c, \tag{6}$$

where the primes denote quantities measured in the frame of the perturbed motion. The mean intensity and radiation flux are defined by

$$J = \frac{1}{4\pi} \int I \, d\Omega; \quad H_i = \frac{1}{4\pi} \int I \mu_i \, d\Omega. \tag{7}$$

Note that in the unperturbed system, there is no net radiation flux, and $4\pi J = \rho_e c^3$

For opacity due solely to Thomson scattering by free electrons,

$$K = n_e \sigma_T, \tag{8}$$

and

$$j' = K' J', \tag{9}$$

where for simplicity we assume that the scattering is isotropic in the rest frame of the electrons. Consequently, one obtains from Equation (6) and (9), to first order in the perturbed quantities,

$$j = KJ(1 + 4 \mu_i v_i/c). \tag{10}$$

The procedure is now to expand the mean intensity in spherical harmonics, bearing in mind that for small perturbations in an anisotropic universe, this provides an adequate representation at arbitrary optical depths. One writes

$$I = I^{(0)} + \mu_i I_i^{(1)} + \left(\frac{3}{2}\mu_i \mu_i - \frac{1}{2}\right) I^{(2)} + \dots, \tag{11}$$

neglecting the higher order terms in the expansion, and takes the first three moments of the transfer Equation (5).

As emphasized by Anderson and Spiegel (1972), it is necessary to proceed to one

higher order in the moments of the transfer equation than in the conventional Eddington approximation in order to obtain the correct solution at large optical depths as given by Thomas (1930). For clarity, I write the moments out explicitly:

$$\frac{\partial H_i}{\partial x_i} + \frac{1}{c} \frac{\partial J}{\partial t} + \frac{1}{c} \frac{4\dot{a}}{a} J = 0; \tag{12}$$

$$\frac{1}{c} \frac{\partial H_i}{\partial t} + \frac{1}{c} \frac{4\dot{a}}{a} H_i + \frac{\partial K_{ij}}{\partial x_j} + K H_i = \frac{4}{3} \frac{v_i}{c} K J; \tag{13}$$

and

$$\frac{1}{c} \frac{\partial K_{ij}}{\partial t} + \frac{1}{c} \frac{4\dot{a}}{a} K_{ij} + \frac{\partial}{\partial x_k} Q_{ijk} + K K_{ij} = \frac{1}{3} K J \delta_{ij}, \tag{14}$$

where

$$K_{ij} = \frac{1}{4\pi} \int I \mu_i \mu_j d\Omega \quad \text{and} \quad Q_{ijk} = \frac{1}{4\pi} \int I \mu_i \mu_j \mu_k d\Omega. \tag{15}$$

The definitions (7) and (15), together with the expansion (11), lead to the consistency relations*

$$J = I^{(0)}; \quad H_i = \frac{1}{3} I_i^{(1)}; \tag{16}$$

$$K_{ij} = \left(\frac{1}{3} I^{(0)} + \frac{2}{15} I^{(2)}\right) \delta_{ij}; \quad Q_{ijk} = \frac{1}{5} I_i^{(1)} \delta_{jk}. \tag{17}$$

Since $I^{(1)} = I^{(2)} = 0$ in the unperturbed system, one can now carry through a perturbation of Equations (12), (13) and (14). The net result is that by combining relations (17) with Equations (12), (13) and (14), one obtains a single equation relating the perturbed mean intensity J_1 to v . After making use of Equation (3), one obtains

$$\left[\left(\frac{\partial}{\partial t} + \frac{4\dot{a}}{a} + Kc \right) \left(\frac{\partial}{\partial t} + \frac{4\dot{a}}{a} + Kc \right) \left(\frac{\partial}{\partial t} + \frac{4\dot{a}}{a} \right) + \frac{2}{3} k^2 c^2 \left(\frac{\partial}{\partial t} + \frac{4\dot{a}}{a} \right) + k^2 \frac{Kc^3}{3} \right] J_1 = \frac{4}{3} \left(\frac{\partial}{\partial t} + \frac{4\dot{a}}{a} + Kc \right) Kc J_s. \tag{18}$$

A Fourier transform has been performed over the spatial dependence of the perturbed quantities, k being the wave number.

One can derive a second equation relating s to the perturbed radiation field from the conservation Equation (4). The radiation stress tensor can be written in the form

$$T_{rad}^{ij} = \frac{4\pi}{c} (J u^i u^j + u^i H^j + u^j H^i + K^{ij}) \tag{19}$$

and the energy-momentum tensor of the matter is

$$T_{mat}^{ij} = -p(g^{ij} + u^i u^j) + \varrho u^i u^j, \tag{20}$$

* Parentheses around indices imply summation over all permutations of the indices, and division by the number of permutations; $\delta_{ij} = 1$ ($i = j$) and is otherwise zero.

where u^i is the velocity four-vector, normalized so that

$$u_i u^i = 1,$$

and the background metric is

$$dS^2 = dt^2 - a^2(dr^2 + r^2 \sin^2 \Theta d\Phi^2 + r^2 d\Theta^2). \tag{21}$$

The conservation Equations (4) yield to first order in the perturbed quantities

$$\frac{\partial v^\lambda}{\partial t} + v^\alpha \frac{\partial v^\lambda}{\partial x_\alpha} = \frac{-1}{\rho} \frac{p_1}{\partial x_\alpha} \delta^{\lambda\alpha} - \frac{c^2}{2} \left(1 + \frac{3}{4}\xi\right) \frac{\partial h_0^0}{\partial x_\alpha} \delta^{\lambda\alpha} - \frac{1}{\rho} \frac{\partial}{\partial x_\alpha} \delta T_{\text{rad}}^{\lambda\alpha}, \tag{22}$$

where

$$\xi \equiv 4\rho_r/3\rho.$$

Taking the divergence of Equation (22), and using Equations (2) and (3) to eliminate h_0^0 and v , one obtains the result

$$\left\{ \frac{\partial^2 s}{\partial t^2} + \left(2 - \frac{\xi}{\frac{4}{3} + \xi}\right) \frac{\dot{a}}{a} \frac{\partial s}{\partial t} \right\} + \frac{k^2 c_m^2}{a^2 (1 + \frac{3}{4}\xi)} \left(s + \frac{T_1}{T}\right) + \frac{4\pi/c k^\alpha k^\beta}{a^2 \rho (1 + \frac{3}{4}\xi)} K_{\alpha\beta} - \frac{1}{\rho (1 + \frac{3}{4}\xi)} \frac{4\pi}{c^2} \left(\frac{\partial}{\partial t} + \frac{2\dot{a}}{a}\right) \frac{ik^\alpha}{a} H_\alpha = 4\pi G \rho \left(s + \frac{3}{4}\xi \frac{J_1}{J}\right). \tag{23}$$

The matter sound velocity c_m and perturbed matter temperature T_1 have been introduced in deriving this result.

In order to specify the matter temperature it is necessary to introduce the energy equation. Only energy transfer by electron scattering (Weymann, 1966) need be considered, and one has

$$\frac{3}{2} \rho k \frac{T}{m_p} (1+x) \frac{\partial}{\partial t} \ln(T \rho^{-2/3}) = 4\rho_r c^2 n_e \sigma_T c k \frac{(T_r - T)}{m_e c^2}, \tag{24}$$

where x is the fractional hydrogen ionization and T_r is the radiation temperature. Carrying through a perturbation of Equation (24), with the requirement that $T = T_r$ in the unperturbed system, leads to the equation

$$\left(\frac{\partial}{\partial t} + \frac{2m_p}{m_e(1+x)} K c \xi\right) \frac{T_r}{T} = \frac{m_p K c \xi}{2m_e(1+x)} \frac{J_1}{J} + \frac{2}{3} \frac{\partial s}{\partial t}. \tag{25}$$

Equations (18), (23), and (25) describe the evolution of matter and radiation perturbations in an Einstein-de Sitter universe at arbitrary optical depths to the radiation field.

3. Adiabatic Fluctuations

The purpose of this section is to summarize for subsequent application the solutions that describe the evolution of adiabatic fluctuations: further details can be found in the review by Field (1971a).

It is convenient to divide the discussion of perturbations into two regimes, corresponding to scales that are either opaque or transparent to the ambient radiation field. I consider fluctuations that are also adiabatic, in the sense that the perturbations are isentropic. In the limit of high optical depth ($k/K \ll 1$), the flux H_α is of higher order in k/K than the mean intensity, and can be neglected in Equation (23). Moreover, the second moment coefficient that enters into the Eddington approximation for $K_{\alpha\beta}$ in Equation (17) can similarly be neglected in Equation (23), which now reduces to

$$\frac{\partial^2 s}{\partial t^2} + \left(2 - \frac{\xi}{\frac{4}{3} + \xi}\right) \frac{\dot{a}}{a} \frac{\partial s}{\partial t} + \frac{k^2}{(1 + \frac{3}{4}\xi)} \frac{c^2 s \xi}{3a^2} = 4\pi G \rho s(1 + \xi). \tag{26}$$

In deriving this equation, use has been made of the optically thick limit of Equation (18),

$$\frac{J_1}{J} = \frac{4s}{3}. \tag{27}$$

I also choose to neglect the fluctuation in the matter temperature: because of the overwhelming contribution of radiation pressure compared to matter pressure, this term never becomes significant prior to the decoupling epoch.

Solutions to Equation (26) can best be studied by considering separately the radiation and matter-dominated epochs. Note that the mass densities of matter and radiation are equal at $\xi = \frac{4}{3}$, or

$$1 + z = 1.04 \times 10^4 \Omega h^2, \tag{28}$$

where $h = (H_0/50 \text{ km s}^{-1} \text{ Mpc}^{-1})$, $\Omega = 8\pi G \rho_0/3H_0^2$, and ρ_0 is the mean density of matter and radiation in the Universe at the present epoch.

3.1. RADIATION-DOMINATED EPOCHS ($\xi \gg 1$)

Equation (26) reduces to

$$\frac{\partial^2 s}{\partial t^2} + \frac{\dot{a}}{a} \frac{\partial s}{\partial t} + \frac{4}{9} \frac{k^2 c^2 s}{a^2} = \frac{16\pi G \rho_0 s}{3}. \tag{29}$$

For a Friedmann universe at $\xi \gg 1$, $a \propto t^{1/2}$, and the solutions to Equation (28) are

$$s \propto t^{1/4} J_{\pm 3/2} \left(\frac{4}{3} \frac{kct}{a} \right). \tag{30}$$

This solution has the expected form (Lifshitz, 1946): if a critical wave number is defined by

$$\frac{k_{cr}}{a} = \frac{3}{4ct}, \tag{31}$$

then for $k \gg k_{cr}$, the solution is a sound wave of constant amplitude, and for $k \ll k_{cr}$, there are two unstable modes, a growing mode $s \propto t$, and a decaying mode, $s \propto t^{-1/2}$. Equation (29) is strictly valid only for wavelengths short compared to the particle

horizon, and is accordingly not accurate in the long wavelength limit at radiation-dominated epochs.

3.2. MATTER-DOMINATED EPOCHS ($\xi \ll 1$)

In general, if γ is the adiabatic index, Equation (26) takes the form

$$\frac{\partial^2 s}{\partial t^2} + \frac{4}{3t} \frac{\partial s}{\partial t} + \left[k^2 c_{s,i}^2 \left(\frac{t_i}{t} \right)^{2(\gamma-1/3)} - \frac{2}{3t^2} \right] s = 0, \tag{32}$$

where we have used the result that $a \propto t^{2/3}$. The epoch t_i is arbitrary, and is introduced in order to remove the explicit dependence on the scale-factor a . $c_{s,i}$ is the sound velocity at t_i . Prior to decoupling of matter and radiation, c_s is given by

$$c_s = \frac{c_\xi^{1/2}}{\sqrt{3}} (1 + \xi)^{-1/2}, \tag{33}$$

and subsequently

$$c_s = c_m = \left(\frac{kT}{\mu m_p} \right)^{1/2}, \tag{34}$$

where μ is the mean molecular weight.

Equation (32) can be solved for $\gamma = \frac{4}{3}$ and $\gamma = \frac{5}{3}$, appropriate to epochs before and subsequent to decoupling.

(i) $\gamma = \frac{4}{3}$.

The solution for s is $s \propto t^{m_\pm}$, where

$$m_\pm = \frac{1}{6} \pm \frac{5}{6} \left(1 - \frac{36}{25} \sigma^2 \right)^{1/2}, \tag{35}$$

and

$$\sigma = k c_{s,i} t_i. \tag{36}$$

σ is identified as the ratio of Jeans length $c_{s,i} t_i$ at epoch t_i to wavelength, and the criteria $\sigma \ll 1$ or $\sigma \gg 1$ divide the solution respectively into an unstable regime, where $s \propto t^{2/3}$ or t^{-1} , and into a damped oscillatory regime, where $s \propto t^{-1/6} \exp(i\sigma \ln t)$.

(ii) $\gamma = \frac{5}{3}$.

In this regime, the perturbations are adiabatic, provided that all coupling with the radiation is neglected. The solutions to Equation (32) are Bessel functions of order $\frac{5}{2}$, that can be written in the explicit form

$$s \propto \left(\frac{3}{x^2} - 1 \right) \sin x - \frac{3}{x} \cos x, \tag{37}$$

and

$$s \propto \left(1 - \frac{3}{x^2} \right) \cos x - \frac{3}{x} \sin x,$$

where

$$x = 3\sigma \left(\frac{t_i}{t} \right)^{1/3}. \quad (38)$$

Here again, there are two regimes to the solution:

$$\begin{aligned} x \gg 1, & \text{ corresponding to damped oscillations, and} \\ x \ll 1, & \text{ corresponding to growth according to } s \propto t^{2/3} \text{ or decay } s \propto t^{-1}. \end{aligned}$$

4. Initial Conditions

The question of appropriate critical conditions for the density perturbations described in the previous section is extremely perplexing. What is self-evident is that the dominant modes increase algebraically with time ($\sim t$ or $\sim t^{2/3}$), and consequently, initial conditions must be specified at some finite time in the past. Extreme difficulties arise in attempting to account for the origin of initial fluctuations of finite amplitude, as already emphasised in Section 1 (cf. also Layzer, 1964).

Consequently, I shall subsequently develop the empirical or phenomenological approach, whereby the presence is postulated of sizable initial fluctuations distributed over all length-scales. One then seeks in various physical processes that occur the dissipation of fluctuations of various scales, and the consequent development of a spectrum that bears some relation to the observed large-scale distribution of matter in the Universe. Various authors have espoused this philosophy, and it is the results of this considerable effort that I intend to review.

First, however, it is important to establish the manner of adiabatic evolution of a spectrum of density perturbations. Formally, one proceeds by postulating appropriate initial conditions on the Fourier-transform of the density perturbation, of the form

$$\begin{aligned} s(\mathbf{k}, t_i) &= s_i(\mathbf{k}, t_i) \\ \dot{s}(\mathbf{k}, t_i) &= \dot{s}_i(\mathbf{k}, t_i). \end{aligned} \quad (39)$$

The general solution for s ,

$$s(\mathbf{k}, t) = A_1 s_1(\mathbf{k}, t) + A_2 s_2(\mathbf{k}, t), \quad (40)$$

contains two arbitrary constants that can now be evaluated, and the required spectrum is given by

$$s(r, t) = \int s(\mathbf{k}, t) e^{i\mathbf{k}\cdot\mathbf{x}} d^3 \mathbf{k}; \mathbf{x} = \mathbf{r}/a \quad (41)$$

Some insight into this procedure can be attained by studying a particular example that allows an analytic treatment. I shall adopt the initial constraints

$$\begin{aligned} s(k, t_i) &= s_0 k^{-2} \\ \dot{s}(k, t_i) &= 0. \end{aligned} \quad (42)$$

These are chosen because of their simplicity: no cut-off in k need be specified for convergence properties. Moreover, the Fourier-transformed initial spectrum has a fairly plausible behavior with scale:

$$s(r, t_i) = 2\pi^2 s_0 r^{-1}. \tag{43}$$

However, it must be emphasized that this is purely an illustrative example.

The transform (41) can be performed exactly for the solution given in the previous section corresponding to radiation-dominated eras prior to decoupling, or to matter-dominated eras subsequent to decoupling. For the $\gamma = \frac{5}{3}$ case, where $s(k, t)$ is given in Equation (37), $s(r, t)$ has the form

$$s(r, t) = \begin{cases} s_0 \frac{\pi^2}{r} \left[2 + \frac{3(y-1)^2}{y} - \frac{2(y-1)^3}{y^3} + \frac{6(y-1)}{5y^3} \right] & \text{for } r > \alpha \frac{y-1}{y} \\ s_0 \frac{\pi^2}{\alpha} \left[\frac{3}{2}(y-1) \left(5 - y - \frac{3}{y} + \frac{1}{y^2} \right) + \frac{r^2}{\alpha^2} - \frac{3}{10} \frac{r^4 y^2}{\alpha^4} \right] & \text{for } r < \alpha \frac{y-1}{y}, \end{cases} \tag{44}$$

where $y = (t/t_i)^{1/3}$ and $\alpha = 3c_s, t_i$.

The asymptotic form of the solution (44) at long times ($y \gg 1$) is

$$s \sim \frac{6}{5} \pi^2 s_0 y^2 r^{-1}. \tag{45}$$

Solutions $s(r, t)$ have been evaluated graphically for expression (44), and are shown in Figure 1.

The relaxation of $s(r, t)$ from the initial perturbation occurs as gravitational effects dominate. At long wavelengths, where pressure effects are negligible, the final state is that of homologous growth as $t^{2/3}$, preserving a spatial dependence of the form (45). Peebles (1971) has given numerical solutions to a related problem.

This result demonstrates the fundamental principle, apparent from inspection of Equation (32), that large wavelength ($k \rightarrow 0$) and/or long-time solutions ($t \rightarrow \infty$) for $\gamma > \frac{4}{3}$ grow at a rate independent of wavelength. The $\gamma = \frac{4}{3}$ solutions show a similar behavior at large wavelengths

$$k/a \ll c_s t. \tag{46}$$

Note that k/a is a comoving wavenumber, and refers to a fixed number of particles. A more realistic spatial dependence for density fluctuations may be obtained by computing the auto-correlation function. For a white-noise power spectrum, the asymptotic dependence is $\langle s(r, t)^2 \rangle^{1/2} \propto r^{-3/2}$.

5. Photon Dissipation and Non-Adiabatic Modes

In order to examine the effects on $s(r, t)$ of dissipation by interaction between matter and radiation, it is necessary to use the full Equation (18) for J_1 derived from radiative transfer theory rather than the adiabatic limit (27).

One can simplify matters somewhat by neglecting higher order radiative terms in

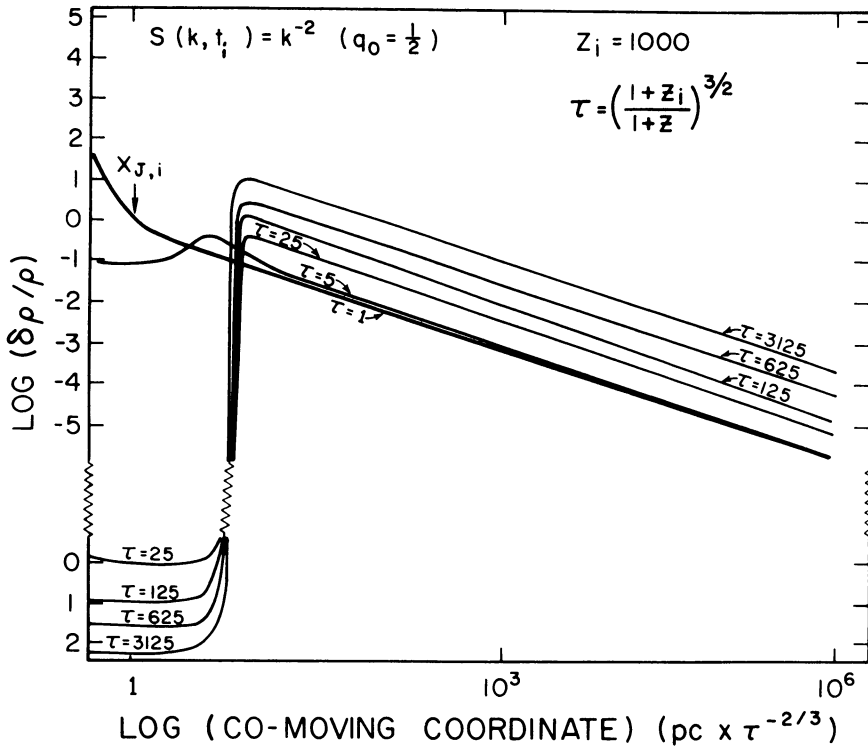


Fig. 1. Density perturbation $\delta\rho/\rho$ as a function of comoving coordinate x at various times $\tau \equiv t/t_i \geq 1$. The assumed initial conditions at $z = 1000$ are defined by Equation (42). The subscript 'J, i' denotes the initial value of the Jeans length. The lower part of the ordinate is labeled to indicate negative values of $\delta\rho/\rho$; the normalization is arbitrary.

Equation (23), which can then be expressed as

$$\frac{\partial^2 s}{\partial t^2} + \left(2 - \frac{\xi}{\frac{4}{3} + \xi}\right) \frac{\dot{a}}{a} \frac{\partial s}{\partial t} + \left(\frac{k^2 c_m^2}{a^2 (1 + \frac{3}{4}\xi)} - 4\pi G\rho\right) s = \left(4\pi G\rho_r - \frac{k^2 c^2}{4a^2} \frac{\xi}{(1 + \frac{3}{4}\xi)}\right) \frac{J_1}{J}. \tag{47}$$

Equations (47) and (18) must now be solved simultaneously to provide the desired solutions. To proceed further, it is instructive to neglect the gravitational terms, and analyze modes of wavelength small compared to ct .

Since in general, one seeks modes that vary on a much more rapid time-scale than the background, one can replace each time derivative by a frequency q that describes the time dependence of the perturbation $\sim e^{qt}$, and ignore all terms of order $(qt)^{-1}$ or higher. One can then obtain a dispersion relation for q and k that can be written in the form

$$(\omega^2 + \omega\xi + 2\alpha^2 d^2) \left[(\omega + 1)^2 \omega + \frac{3}{5}\alpha^2 \omega + \frac{\alpha^2}{3} \right] = \xi \omega^2 (\omega + 1). \tag{48}$$

Dimensionless parameters have been introduced, defined by

$$\omega = \frac{q}{KC}; \quad \alpha = \frac{k}{aK}; \quad \text{and} \quad d = \frac{c_m}{c(1 + \frac{3}{4}\xi)^{1/2}} \tag{49}$$

The scattering coefficient K is identified as

$$K = n_e \sigma_T; \tag{50}$$

consequently the dimensionless wavenumber α is inversely proportional to the optical depth across a fluctuation.

At this point, it is convenient to divide our discussion into two regimes, corresponding to opaque ($\alpha \ll 1$) or transparent ($\alpha \gg 1$) fluctuations.

5.1. OPAQUE FLUCTUATIONS ($\alpha \ll 1$)

This regime has been treated for a static medium by Field (1971b), whose dispersion relation in this limit is almost identical to Equation (48). The modes can be analyzed by expanding ω in powers of α : note that $\omega/\alpha = q/kc \equiv qt(kct)^{-1}$, and will be small compared with unity for wavelength satisfying $kct \gg 1$. The result of this calculation is that two principal damping modes appear:

$$\omega_1 = \frac{-\alpha^2 d^2}{\xi + 3d^2} \tag{51}$$

and

$$\omega_2 = \frac{i\alpha}{\sqrt{3}} \left(\frac{\xi + 3d^2}{1 + \xi} \right)^{1/2} - \frac{\alpha^2}{6(1 + \xi)^2} \left\{ 1 + \frac{4\xi}{5}(1 + \xi) - \frac{3(1 + 2\xi - 3\xi d^2)}{\xi + 3d^2} d^2 \right\}. \tag{52}$$

Mode ω_1 corresponds to the damped isothermal mode of Zel'dovich (1966) for a matter fluctuation in a uniform radiation field. One can see this directly by applying the moment Equations (12) and (13) to Equation (23), and taking the isothermal limit $J_1 \rightarrow 0$. The resulting equation is

$$\frac{\partial^2 s}{\partial t^2} + \left(\frac{2\dot{a}}{a} - \frac{\xi}{(\frac{4}{3} + \xi)} \frac{\dot{a}}{a} + \frac{Kc\xi}{1 + \frac{3}{4}\xi} \right) \frac{\partial s}{\partial t} + \left(\frac{k^2 c_m^2}{a^2 (1 + \frac{3}{4}\xi)} - 4\pi G\rho \right) s = 0. \tag{53}$$

Upon applying a similar analysis to that given above to the general case, one obtains the result that the principal mode of Equation (53) is

$$\omega = -\alpha^2 d^2 / \xi. \tag{54}$$

The net effect of the damping of the isothermal mode is that the amplitude is frozen prior to decoupling, the self-gravity of the perturbation being in balance with the radiation drag force.

The second mode (52) is an acoustic wave, propagating at the adiabatic sound velocity c_s , and damped by photon diffusion. The existence of this damping was established by Silk (1967, 1968), and independently in unpublished work by Michie

(1967) and Peebles (1967). The general expression for the damping rate (52) differs slightly, but not significantly, from that given by Field (1971b): the rates agree if $d=0$ for arbitrary ξ , or if $\xi \ll 1$. Note that expression (52) is applicable for arbitrary ratios of matter to radiation.

5.2. TRANSPARENT FLUCTUATIONS ($\alpha \gg 1$)

Consider next the limiting solutions to the dispersion relation (48) at large α . One can readily establish that, when

$$\alpha^2 \gg \xi, \tag{55}$$

the principal modes are given by

$$u = -\frac{3}{5} \frac{\alpha^2 \xi}{1 + \frac{3}{5} \alpha^2} \quad \text{or} \quad -\frac{\alpha^2}{1 + \frac{3}{5} \alpha^2}. \tag{56}$$

The modes in Equation (56) are due to damping by radiation drag, occurring predominantly at an optical depth of order unity through the fluctuation.

6. Damping During the Decoupling Epoch

Since the preceding analysis is valid in general for arbitrary optical depths, it is a straightforward matter to evaluate the damping rates before, during and after the decoupling era at $z \sim 1000$. Although Equations (18) and (47) can in principle be solved numerically, it is adequate for the present purpose to use the dispersion relation (48) in order to study the evolution of density perturbations over comoving length-scales short compared to the particle horizon. One can therefore apply the asymptotic forms of the dispersion relation (52) and (46), and interpolate between them as necessary. The results of such a calculation are shown in Figure 2 where the exponential damping factor γ is plotted as a function of time through decoupling for various comoving mass-scales of interest in two different cosmological models, corresponding to a spatially flat ($\Omega = 1$) and to a low density ($\Omega = 0.02$) universe. The definition of γ is such that the amplitude of an initially adiabatic density perturbation is damped by a factor $\exp \gamma$ at any given epoch. Consequently the maximum value of γ for any specified mass-scale gives the final damping for that scale.

The mass-scales chosen are intended to span the range of significant damping. The comoving mass-scale is defined by

$$M = \frac{\pi}{6} \rho \left(\frac{2\pi a}{k} \right)^3. \tag{57}$$

Numerically, one obtains

$$\lambda \equiv 2\pi a/k = \frac{4.35 \times 10^{24}}{1+z} M_{11}^{1/3} (\Omega h^2)^{-1/3} \text{ cm}, \tag{58}$$

where $M_{11} = M/10^{11} M_{\odot}$.

The main results are consistent with those found by Peebles and Yu (1970). Mass-scales of up to $10^{12} M_{\odot}$ are severely damped in an $\Omega = 1$ universe, and up to $10^{15} M_{\odot}$ if $\Omega = 0.02$. The Hubble constant H_0 has been set equal to $50 \text{ km (s Mpc)}^{-1}$. The run of ionization used in the present calculation is taken from the work of Peebles (1968) and Zel'dovich *et al.* (1969), and is shown in Figure 2.

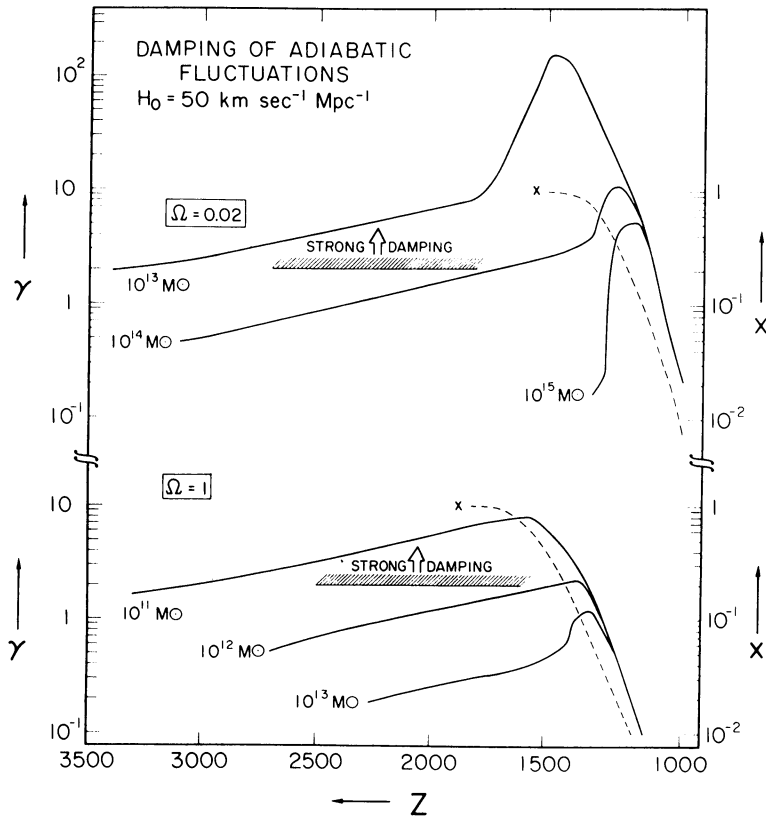


Fig. 2. Exponential damping factor γ evaluated as a function of time for several comoving mass scales, in $\Omega = 1$ and $\Omega = 0.02$ cosmological models. The variation of $x(t)$ is also shown as a dashed line for each model.

It is apparent that the bulk of the damping occurs when the hydrogen ionization level has already fallen considerably. As previously indicated, the relatively dominant peaks in the damping rates are primarily due to radiation drag when the optical depth across the fluctuation is of order unity.

To examine further the physical basis of this mechanism, one can define the optical depth over scale λ ,

$$\tau_{\lambda} = n_e \sigma_T \lambda = 6.9 \times 10^{-6} (1+z)^2 \times M_{11}^{1/3} (\Omega h^2)^{2/3}. \tag{59}$$

The epoch at which $\tau_\lambda = 1$ can be expressed in the form *

$$z|_{\tau_\lambda=1} = \frac{4.37 \times 10^4}{8.3 + \ln \left(\frac{M_{11} z^3}{h\Omega^{1/2}} \right)}, \tag{60}$$

and it is apparent from Figure 2 that the peak damping occurs in the vicinity of $z|_{\tau_\lambda=1}$.

An important feature of the solutions exhibited in Figure 2 is that radiation drag effectively damps much of the amplitude jump that adiabatic fluctuations initially below the Jeans mass would otherwise experience at decoupling, apparently contrary to the assertion of Sunyaev and Zeldovich (1970). One can investigate this result in more detail by considering the amplitude of a fluctuation that is Jeans-stable prior to decoupling, but is subsequently Jeans-unstable when the sound velocity has dropped from c_s to c_m . Denoting quantities evaluated immediately prior to decoupling with a minus sign and subsequent to decoupling with a plus sign, one has from the perturbed continuity equation that

$$\frac{v_+}{v_-} = \frac{1}{3} \frac{\lambda}{c_{s-}} \frac{\dot{a} s_+}{a s_-} = 1.36 \times 10^{-4} \left(1 + \frac{7.8 \cdot 10^3 \Omega h^2}{z} \right)^{1/2} z^{1/2} M_{11}^{1/3} (\Omega h^2)^{1/6} \left(\frac{s_+}{s_-} \right). \tag{62}$$

If, as Sunyaev and Zel'dovich argue, there can be no discontinuity in velocity, one would therefore have to have an amplitude increase amounting to a factor

$$s_+/s_- = 14 M_{11}^{-1/3} \text{ for } \Omega = 1 \text{ or } s_+/s_- = 380 M_{11}^{-1/3} \text{ for } \Omega = 0.02.$$

In fact, as has already been shown, radiation drag does produce a severe reduction in velocity: only at large mass-scales is it effective, and here the magnitude of a possible amplitude jump is proportionately diminished according to (62).

To further establish this result, one can define a characteristic time-scale for radiation drag,

$$t_{\text{drag}} = \frac{m_p c}{\sigma_T \rho_e c^2 x} = \frac{1.85 \times 10^{23} \text{ s}}{x(1+z)^4}. \tag{63}$$

At $z|_{\tau_\lambda=1}$ appropriate for $M = 10^{12} M_\odot$ ($z = 1350$), one has $t_{\text{drag}} = 3.3 \times 10^{11} \text{ s}$. An appropriate time for comparison is that fraction of an expansion time Δt_{exp} evaluated at $z|_{\tau_\lambda=1}$ over which x does not decrease by an appreciable factor. One can estimate such a time-scale by defining

$$\Delta t_{\text{exp}} = \frac{1}{2} t_{\text{exp}} \left(\frac{d \ln x}{d \ln t} \right)^{-1}, \tag{64}$$

where $t_{\text{exp}} = 4.1 \times 10^{17} (\Omega h^2)^{-1/2} z^{-3/2} \text{ s}$ at $z \gg 1$.

One obtains $\Delta t_{\text{exp}}(z = 1350) = 6.3 \times 10^{11} \text{ s}$, if $\Omega = h = 1$.

* In deriving this expression, a simple analytical expression is used for the fractional ionization,

$$x = \frac{1.1 \times 10^7}{z \Omega^{1/2} h} \exp(-1.458 \times 10^4/z), \tag{61}$$

which is valid over $1500 \gtrsim z \gtrsim 900$ (Sunyaev and Zel'dovich, 1970).

An estimate of the predicted damping by radiation drag is given by setting $\gamma = \Delta t_{\text{exp}}/t_{\text{drag}}$, and one obtains a damping factor in agreement with the results shown in Figure 1 for $M = 10^{12} M_{\odot}$, $\Omega = 1$. Note that the amplitude decrease amounts to a factor 7, which is just adequate to reduce v_+/v_- to unity and entirely remove any amplitude jump effect that could otherwise occur.

Similarly for a perturbation containing a mass $M = 10^{14} M_{\odot}$ in an $\Omega = 0.02$ universe, one obtains $z|_{\tau_{\lambda=1}} = 1140$ and $\Delta t_{\text{exp}}/t_d = 8.3$, again consistent with Figure 2. More generally, a simple analytic expression for the damping factor γ can be derived by utilizing expression (61) for $x(t)$. One obtains

$$\gamma = 8.4 \times 10^4 \beta^{-5/2} (\Omega h^2)^{-1} \exp(-10.8\beta), \tag{65}$$

where

$$\beta = 1 + 0.03 \ln [M_{12} (\Omega h^2)^{-1/2}] \tag{66}$$

and

$$M_{12} = M/10^{12} M_{\odot}.$$

The velocity overshoot effect does give rise to a density amplitude jump for fluctuations of larger mass; however the effect is necessarily small in view of the mass dependence of Equation (62).

7. Conclusions

The principal aim of this paper has been to establish the spectrum of primordial adiabatic density fluctuations. Damping before and during decoupling imposes a lower bound on the surviving mass-scales, of from $M_D \equiv 10^{12}$ to $10^{15} M_{\odot}$ for $\Omega = 1$ to 0.01 respectively, with $h = 1$. The critical mass-scale for damping M_D is defined by

$$\gamma_{\text{max}}(M_D) = 1.$$

The preceding results on adiabatic fluctuations can be combined in a single expression that approximately represents the spectrum of primordial adiabatic density fluctuations immediately after decoupling. For any mass-scale M , this spectrum, and its subsequent evolution at epochs $z > \Omega^{-1}$, is given by

$$\langle (\delta\varrho/\varrho)^2 \rangle^{1/2} \approx 3 \times 10^{-3} (t/t_d)^{2/3} (M_D/M)^{1/2} \exp(-(M_D/M)^{2/3}).$$

The asymptotic behaviour of this spectrum is determined at short wavelengths by dissipative processes, and at long wavelengths by the adoption of a white noise power spectrum for the initial distribution of density fluctuations. This latter assumption can be justified in a phenomenological way because of the success one consequently achieves when the resulting spectrum of gravitationally bound systems is compared with observational data on the distribution of galaxies (Balko, 1971; Peebles, 1973).

Such a spectrum would suffice to generate the formation of bound systems on the

scale of galaxies or small galaxy clusters by $z \gtrsim 1$ if $\Omega \sim 0.1-1$. The initial conditions required for this to occur are $|\delta\rho/\rho|_i \sim 10^{-3}$ at $z_i \sim 10^6$. Extrapolation of the initial conditions to earlier epochs requires a fully general relativistic treatment and the resulting time-dependence of $\delta\rho/\rho$ on scales $\gtrsim ct$ is coordinate gauge-dependent.

Two other characteristic mass-scales emerge from the theory that has been described here. Scales larger than the Jeans mass immediately prior to decoupling have never passed through an oscillatory phase, and have undergone uninterrupted growth (Field and Shepley, 1968). The minimum mass for this to occur is about $5 \times 10^{16} M_\odot$ (Peebles and Yu, 1970).

Isothermal fluctuations remain frozen in until decoupling, when they become unstable to the growing mode if above the Jeans mass M_J . After decoupling

$$M_J = 2.9 \times 10^5 (T/T_D)^{3/2} (\Omega h^2)^{-1/2} M_\odot,$$

where T is the matter temperature, and $T_D \sim 4000$ K is its value at decoupling.

On the other hand, prior to decoupling, the radiation pressure maintains a high sound speed c_s , and

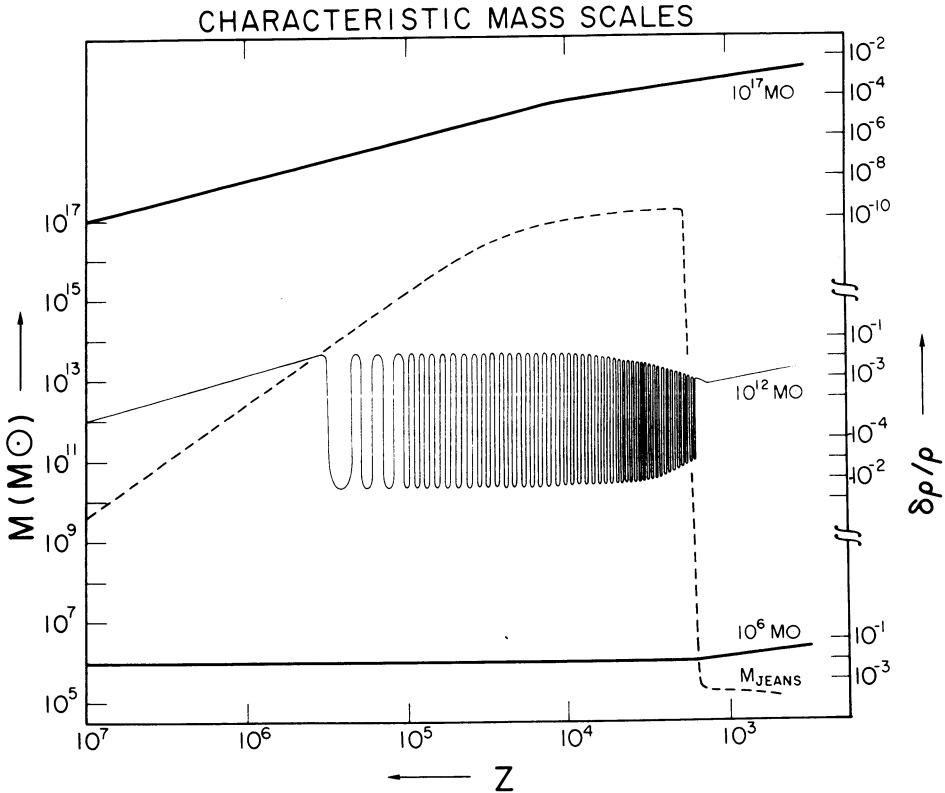


Fig. 3. The evolution of the characteristic mass-scales of density perturbations in an Einstein-de Sitter universe. The variation of the Jeans mass with redshift is shown as a dashed line.

$$M_J = 2 \times 10^{23} (\Omega h^2)^{-1/2} (1+z)^{-3/2} \xi^{3/2} (1+\xi)^{-3} M_\odot,$$

where

$$\xi \equiv \frac{4q_r}{3q} = 1.29 \times 10^{-4} (\Omega h^2)^{-1} (1+z). \quad (69)$$

The radiation temperature has been set equal to 2.7 K at the present epoch. Equations (67) and (68) give the two additional critical masses, if evaluated at decoupling.

These results are summarized in Figure 3, where the evolution is depicted of three characteristic mass-scales in an $\Omega = 1$ universe. The amplitude scale has been normalized to allow formation of bound systems prior to the present epoch.

The question remains unanswered of the origin of the assumed initial conditions. However the fact that galaxies, which appear to be a basic primordial constituent of the Universe, roughly coincide in mass with the damping limit (M_D) in a dense universe is perhaps the most encouraging result to have emerged from the theory. The formation of dwarf galaxies can presumably be accounted for by non-linear interactions and subsequent fragmentation of larger systems.

Some additional remarks can be made about consequences of assuming finite amplitude perturbations. This must necessarily involve the break-down of the linear theory, as evidenced in all probability by the formation of shocks. There are at least two ways in which shock formation can be envisaged. An acoustic wave of amplitude S steepens into a shock after S^{-1} periods. This result would be most relevant for mass-scales of interest prior to decoupling. On the other hand, a growing mode of wavelength will develop into a shock over length scales

$$\lambda \gtrsim c_s t S^{-1} \equiv \lambda_j S^{-1}, \quad (70)$$

where its motion becomes supersonic. This effect could be significant after decoupling.

Shock formation offers the possibility of a considerable amplification factor, since once the non-linear regime is entered, gravitational collapse can readily be initiated. Shocks can also provide a mechanism for generating entropy fluctuations from initially purely adiabatic perturbations, as does also the occurrence of dissipation of acoustic waves by radiation damping.

Acknowledgements

I wish to thank Drs J. Bardeen, G. B. Field, and P. J. E. Peebles for stimulating discussions on topics relevant to this research. It is a pleasure to thank Susan Lea for preparing Figure 1.

This work has been supported in part by NASA grant NGR 05-003-453.

References

- Anderson, J. L. and Spiegel, E. A.: 1972, *Astrophys. J.* **171**, 127.
 Balko, A. V.: 1971, *Soviet Phys. JETP* **33**, 237.
 Bardeen, J.: 1968, *Astron. J.* **73**, S164.

- Dautcourt, G. and Wallis, G.: 1968, *Fortschr. Physik* **16**, 545.
- Field, G. B.: 1971a, *Stars and Stellar Systems* **9**, in press.
- Field, G. B.: 1971b, *Astrophys. J.* **165**, 29.
- Field, G. B. and Shepley, L. C.: 1968, *Astrophys. Space Sci.* **1**, 309.
- Irvine, W. M.: 1965, *Ann. Phys., N.Y.* **32**, 322.
- Lanczos, C.: 1925, *Z. Phys.* **31**, 112.
- Layzer, D.: 1964, *Ann. Rev. Astron., Astrophys.* **2**, 341.
- Layzer, D.: 1968, in M. Chretien, S. Deser, and J. Goldstein (eds.), *Astrophysics and General Relativity* **2**, New York, Gordon and Breach, p. 15.
- Lifshifz, E.: 1946, *J. Phys. USSR* **10**, 116.
- Michie, R. W.: 1967, private communication.
- Misner, C. W.: 1967, *Nature* **216**, 40.
- Peebles, P. J. E.: 1967, *Proc. 4th Conf. Relativistic Astrophysics*, New York.
- Peebles, P. J. E.: 1968, *Astrophys. J.* **153**, 1.
- Peebles, P. J. E.: 1969, *Astrophys. J.* **157**, 1075.
- Peebles, P. J. E.: 1973, in preparation.
- Peebles, P. J. E. and Yu, J. T.: 1970, *Astrophys. J.* **162**, 815.
- Rees, M. J.: 1972, *Phys. Rev. Letters* **28**, 1669.
- Silk, J.: 1966, *Astrophys. J.* **143**, 689.
- Silk, J.: 1967, *Nature* **215**, 1155.
- Silk, J.: 1968, *Astrophys. J.* **151**, 459.
- Sunyaev, R. A. and Zel'dovich, Ya. B.: 1970, *Astrophys. Space Sci.* **7**, 1.
- Thomas, L. H.: 1930, *Quart. J. Math., Oxford* **1**, 239.
- Unno, W. and Spiegel, E. A.: 1966, *Publ. Astron. Soc. Japan* **18**, 85.
- Weymann, R. J.: 1966, *Astrophys. J.* **145**, 560.
- Zel'dovich, Ya. B.: 1966, *Uspekhi Fiz. Nauk* **89**, 674 [translated in *Soviet Phys.-Uspekhi* **9**, 602 (1967)].
- Zel'dovich, Ya. B., Kurt, V. G., and Sunyaev, R. A.: 1969, *Soviet Phys. JETP* **28**, 146.

DISCUSSION

Ozernoy: A similar analytical method of investigating the damping of acoustic perturbations during and after the decoupling epoch was elaborated by G. V. Chibisov of our group at the Lebedev Institute of Physics (see *Astron. Zh.* **49**, 74, 1972). His results are very similar to those obtained independently by Dr Silk. One difference may be mentioned. For $\Omega = 1$ the damped mass, according to Chibisov, is approximately one order of magnitude greater than that given by Silk. This difference is significant for the problem of whether the isolation of galaxies occurs independently of the formation of clusters of galaxies or whether it is closely related to them. The isolation of protogalaxies from reasonable inhomogeneities may occur independently of the formation of clusters of galaxies (which could form by means of clustering) only if M_D is as small as $10^{12} M_\odot$. However, if M_D is of the order of or greater than $10^{13} M_\odot$, then the picture of the birth of galaxies must be drastically different: galaxies will form more or less simultaneously with the isolation of protoclusters by means of the fragmentation of the latter.

Silk: The analytical results of Chibisov appear to be based on an over-simplification of the relevant physics involved in matter-radiation interaction during decoupling, and are in serious disagreement with the calculations by Peebles and Yu and myself for $\Omega \gtrsim 0.1$.

Zel'dovich: The mass $M_D \approx \rho \lambda_D^3$ where λ_D is the dissipation scale, determines the minimum mass of fluctuations which survive the radiation dominated era but it is not necessarily the minimum mass for galaxies. The non-linear theory of the origin of galaxies (see the contribution by Doroshkevich, Zel'dovich and Sunyaev) permits much smaller masses because this theory contains characteristic scales much smaller than λ_D .

Silk: The theory that I have described is a linear theory. Provided that the amplitude of the density fluctuations remains sufficiently small, non-linear effects are unimportant. For example, a co-moving mass scale of $10^{12} M_\odot$ must have amplitude $|\delta\rho/\rho| \lesssim 10^{-2}$ in the oscillatory phase prior to decoupling.