Fundamentals of Cosmology

(3) The Formation of Structure in the Universe

- Jeans' Instability in the Expanding Universe
- Non-relativistic case
- Peculiar and Rotational Velocities
- Relativistic case
- The Basic Problem of Structure Formation
- Aspects of the Thermal History of the Universe
- The Standard Λ-Cold Dark Matter Picture
- Non-linear Development of Perturbations

The Object of the Exercise

The aim of the cosmologist is to explain how large-scale structures formed in the expanding Universe in the sense that, if $\delta\varrho$ is the enhancement in density of some region over the average background density ϱ, the *density contrast* $\Delta = \delta\varrho/\varrho$ reached amplitude 1 from initial conditions which must have been remarkably isotropic and homogeneous. Once the initial perturbations have grown in amplitude to $\Delta = \delta\varrho/\varrho \approx 1$, their growth becomes non-linear and they rapidly evolve towards bound structures in which star formation and other astrophysical process lead to the formation of galaxies and clusters of galaxies as we know them.

The density contrasts $\Delta = \delta\varrho/\varrho$ for galaxies, clusters of galaxies and superclusters at the present day are about $\sim 10^6,\ 1000$ and a few respectively. Since the average density of matter in the Universe ϱ changes as $(1+z)^3$, it follows that typical galaxies must have had $\Delta = \delta\varrho/\varrho \approx 1$ at a redshift $z \approx 100$. The same argument applied to clusters and superclusters suggests that they could not have separated out from the expanding background at redshifts greater than $z \sim 10$ and 1 respectively.

The Wave Equation for the Growth of Small Density Perturbations (1)

The standard equations of gas dynamics for a fluid in a gravitational field consist of three partial differential equations which describe (i) the conservation of mass, or the equation of continuity, (ii) the equation of motion for an element of the fluid, Euler's equation, and (iii) the equation for the gravitational potential, Poisson's equation.

$$\text{Equation of Continuity} \quad : \quad \frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \boldsymbol{v}) = 0\,; \tag{1}$$

$$\text{Equation of Motion} \quad : \quad \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\varrho}\nabla p - \nabla \phi\,; \tag{2}$$

$$\text{Gravitational Potential} \quad : \quad \nabla^2 \phi = 4\pi G \varrho\,. \tag{3}$$

These equations describe the dynamics of a fluid of density ϱ and pressure p in which the velocity distribution is $\boldsymbol{v}$. The gravitational potential ϕ at any point is given by Poisson's equation in terms of the density distribution ϱ.

The partial derivatives describe the variations of these quantities *at a fixed point in space*. This coordinate system is often referred to as *Eulerian coordinates*.

The Wave Equation for the Growth of Small Density Perturbations (2)

We need to go through a slightly complex procedure to derive the second-order differential equation. We need to convert the expressions into Lagrangian coordinates, which follow the motion of an element of the fluid:

$$\begin{aligned} \frac{\mathrm{d}\varrho}{\mathrm{d}t} &= -\varrho\nabla\cdot\boldsymbol{v}\ ; && (4) \\ \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} &= -\frac{1}{\varrho}\nabla p - \nabla\phi\ ; && (5) \\ \nabla^2\phi &= 4\pi G\varrho\ . && (6) \end{aligned}$$

Next, we need to put in the uniform expansion of the unperturbed density distribution $\boldsymbol{v} = H_0\boldsymbol{r}$. The unperturbed solutions are then

$$\begin{aligned} \frac{\mathrm{d}\varrho_0}{\mathrm{d}\,t} &= -\varrho_0\nabla\cdot\boldsymbol{v}_0\ ; && (7) \\ \frac{\mathrm{d}\boldsymbol{v}_0}{\mathrm{d}t} &= -\frac{1}{\varrho_0}\nabla p_0 - \nabla\phi_0\ ; && (8) \\ \nabla^2\phi_0 &= 4\pi G\varrho_0\ . && (9) \end{aligned}$$

The Wave Equation for the Growth of Small Density Perturbations (3)

Then, we perturb the system about the uniform expansion $\boldsymbol{v} = H_0 \boldsymbol{r}$:

$$\boldsymbol{v} = \boldsymbol{v}_0 + \delta \boldsymbol{v}, \quad \varrho = \varrho_0 + \delta \varrho, \quad p = p_0 + \delta p, \quad \phi = \phi_0 + \delta \phi \, . \tag{10}$$

After a bit of algebra, we find the following equation for adiabatic density perturbations $\Delta = \delta\varrho/\varrho_0$:

$$\frac{\mathrm{d}^2 \Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = \frac{c_\mathrm{s}^2}{\varrho_0 a^2} \nabla_\mathrm{c}^2 \delta\varrho + 4\pi G \delta\varrho \, . \tag{11}$$

where the adiabatic sound speed c_s^2 is given by $\partial p / \partial \varrho = c_\mathrm{s}^2$. We now seek wave solutions for Δ of the form $\Delta \propto \exp \mathrm{i}(\boldsymbol{k}_\mathrm{c} \cdot \boldsymbol{r} - \omega t)$ and hence derive a wave equation for Δ.

$$\boxed{\frac{\mathrm{d}^2 \Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = \Delta(4\pi G \varrho_0 - k^2 c_\mathrm{s}^2) \, ,} \tag{12}$$

where $\boldsymbol{k}_c$ is the wavevector in comoving coordinates and the proper wavevector $\boldsymbol{k}$ is related to $\boldsymbol{k}_c$ by $\boldsymbol{k}_c = a\boldsymbol{k}$. This is a key equation we have been seeking.

The Jeans' Instability (1)

The differential equation for gravitational instability in a static medium is obtained by setting $\dot{a} = 0$. Then, for waves of the form $\Delta = \Delta_0 \; \exp \mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)$, the dispersion relation,

$$\omega^2 = c_\mathrm{s}^2 k^2 - 4\pi G \varrho_0 \; , \tag{13}$$

is obtained.

- If $c_\mathrm{s}^2 k^2 > 4\pi G \varrho_0$, the right-hand side is positive and the perturbations are oscillatory, that is, they are sound waves in which the pressure gradient is sufficient to provide support for the region. Writing the inequality in terms of wavelength, stable oscillations are found for wavelengths less than the critical *Jeans' wavelength* λ_J

$$\lambda_\mathrm{J} = \frac{2\pi}{k_\mathrm{J}} = c_\mathrm{s} \left(\frac{\pi}{G\varrho} \right)^{1/2} \; . \tag{14}$$

The Jeans' Instability (2)

- If $c_\mathrm{s}^2 k^2 < 4\pi G\varrho_0$, the right-hand side of the dispersion relation is negative, corresponding to unstable modes. The solutions can be written

$$\Delta = \Delta_0 \exp(\Gamma t + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) \,, \tag{15}$$

where

$$\Gamma = \pm\left[4\pi G\varrho_0\left(1 - \frac{\lambda_\mathrm{J}^2}{\lambda^2}\right)\right]^{1/2} \,. \tag{16}$$

The positive solution corresponds to exponentially growing modes. For wavelengths much greater than the Jeans' wavelength, $\lambda \gg \lambda_\mathrm{J}$, the growth rate Γ becomes $(4\pi G\varrho_0)^{1/2}$. In this case, the characteristic growth time for the instability is

$$\tau = \Gamma^{-1} = (4\pi G\varrho_0)^{-1/2} \sim (G\varrho_0)^{-1/2} \,. \tag{17}$$

This is the famous *Jeans' Instability* and the time scale τ is the typical collapse time for a region of density ϱ_0.

The Jeans' Instability (3)

The physics of this result is very simple. The instability is driven by the self-gravity of the region and the tendency to collapse is resisted by the internal pressure gradient. Consider the pressure support of a region with pressure p, density ϱ and radius r. The equation of hydrostatic support for the region is

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{G\varrho M(< r)}{r^2}\,. \tag{18}$$

The region becomes unstable when the self-gravity of the region on the right-hand side of (18) overwhelms the pressure forces on the left-hand side. To order of magnitude, we can write $\mathrm{d}p/\mathrm{d}r \sim -p/r$ and $M \sim \varrho r^3$. Therefore, since $c_\mathrm{s}^2 \sim p/\varrho$, the region becomes unstable if $r > r_\mathrm{J} \sim c_\mathrm{s}/(G\varrho)^{1/2}$. Thus, the Jeans' length is the scale which is just stable against gravitational collapse.

Notice that the expression for the Jeans' length is just the distance a sound wave travels in a collapse time.

The Jeans' Instability in an Expanding Medium

We return first to the full version of the differential equation for Δ.

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = \Delta(4\pi G\varrho - k^2 c_s^2)\ . \tag{19}$$

The second term $2(\dot{a}/a)(\mathrm{d}\Delta/\mathrm{d}t)$ modifies the classical Jeans' analysis in crucial ways. It is apparent from the right-hand side of (19) that the Jeans' instability criterion applies in this case also but the growth rate is significantly modified. Let us work out the growth rate of the instability in the long wavelength limit $\lambda \gg \lambda_J$, in which case we can neglect the pressure term $c_s^2 k^2$. We therefore have to solve the equation

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = 4\pi G\varrho_0\Delta\ . \tag{20}$$

Before considering the general solution, let us first consider the special cases $\Omega_0 = 1$ and $\Omega_0 = 0$ for which the scale factor-cosmic time relations are $a = (\frac{3}{2}H_0 t)^{2/3}$ and $a = H_0 t$ respectively.

The Jeans' Instability in an Expanding Medium

The Einstein–de Sitter Critical Model $\Omega_0 = 1$. In this case,

$$4\pi G\varrho = \frac{2}{3t^2} \qquad \text{and} \qquad \frac{\dot{a}}{a} = \frac{2}{3t} \, . \tag{21}$$

Therefore,

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + \frac{4}{3t}\frac{\mathrm{d}\Delta}{\mathrm{d}t} - \frac{2}{3t^2}\Delta = 0 \, . \tag{22}$$

By inspection, it can be seen that there must exist power-law solutions of (22) and so we seek solutions of the form $\Delta = at^n$. Hence

$$n(n-1) + \tfrac{4}{3}n - \tfrac{2}{3} = 0 \, , \tag{23}$$

which has solutions $n = 2/3$ and $n = -1$. The latter solution corresponds to a decaying mode. The $n = 2/3$ solution corresponds to the growing mode we are seeking, $\Delta \propto t^{2/3} \propto a = (1+z)^{-1}$. This is the key result

$$\boxed{\Delta = \frac{\delta\varrho}{\varrho} \propto (1+z)^{-1} \, .} \tag{24}$$

In contrast to the *exponential* growth found in the static case, the growth of the perturbation in the case of the critical Einstein–de Sitter universe is *algebraic*.

The Jeans' Instability in an Expanding Medium

The Empty, Milne Model $\Omega_0 = 0$ In this case,

$$\varrho = 0 \qquad \text{and} \qquad \frac{\dot{a}}{a} = \frac{1}{t}\,, \tag{25}$$

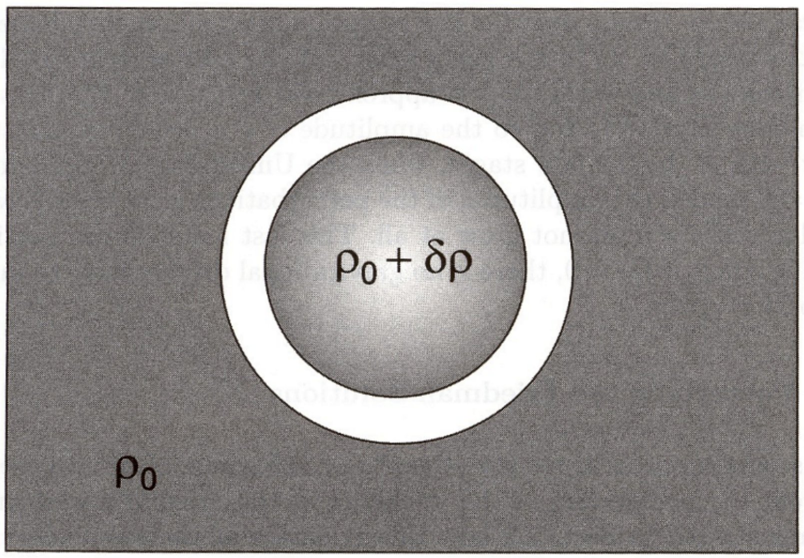
and hence

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + \frac{2}{t}\frac{\mathrm{d}\Delta}{\mathrm{d}t} = 0\,. \tag{26}$$

Again, seeking power-law solutions of the form $\Delta = at^n$, we find $n = 0$ and $n = -1$, that is, there is a decaying mode and one of constant amplitude Δ = constant.

These simple results describe the evolution of small amplitude perturbations, $\Delta = \delta\varrho/\varrho \ll 1$. In the early stages of the matter-dominated phase, the dynamics of the world models approximate to those of the Einstein–de Sitter model, $a \propto t^{2/3}$, and so the amplitude of the density contrast grows linearly with a. In the late stages at redshifts $\Omega_0 z \ll 1$, when the Universe may approximate to the $\Omega_0 = 0$ model, the amplitudes of the perturbations grow very slowly and, in the limit $\Omega_0 = 0$, do not grow at all.

Perturbing the Friedman solutions



Let us derive the same results from the dynamics of the Friedman solutions. The development of a spherical perturbation in the expanding Universe can be modelled by embedding a spherical region of density $\varrho + \delta\varrho$ in an otherwise uniform Universe of density ϱ. The parametric solutions for the dynamics of the world models can be written

$$a = A(1 - \cos\theta) \qquad t = B(\theta - \sin\theta)\ ;$$
$$A = \frac{\Omega_0}{2(\Omega_0 - 1)} \qquad B = \frac{\Omega_0}{2H_0(\Omega_0 - 1)^{3/2}}\ .$$

Perturbing the Friedman solutions

We now compare the dynamics of the region of slightly greater density with that of the background model. We expand the expressions for a and t to fifth order in θ. The solution is

$$a = \Omega_0^{1/3} \left(\frac{3H_0 t}{2}\right)^{2/3} \left[1 - \frac{1}{20}\left(\frac{6t}{B}\right)^{2/3}\right] . \tag{27}$$

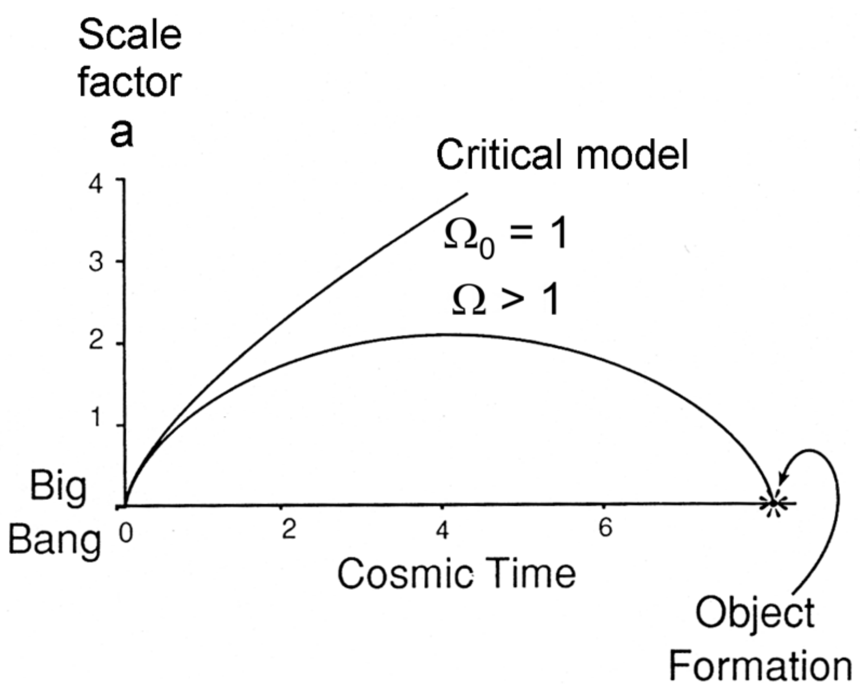
We can now write down an expression for the change of density of the spherical perturbation with cosmic epoch

$$\varrho(a) = \varrho_0 a^{-3} \left[1 + \frac{3}{5}\frac{(\Omega_0 - 1)}{\Omega_0} a\right] . \tag{28}$$

Notice that, if $\Omega_0 = 1$, there is no growth of the perturbation. The density perturbation may be considered to be a mini-Universe of slightly higher density than $\Omega_0 = 1$ embedded in an $\Omega_0 = 1$ model. Therefore, the density contrast changes with scale factor as

$$\Delta = \frac{\delta\varrho}{\varrho} = \frac{\varrho(a) - \varrho_0(a)}{\varrho_0(a)} = \frac{3}{5}\frac{(\Omega_0 - 1)}{\Omega_0} a . \tag{29}$$

Perturbing the Friedman solutions



This result indicates why density perturbations grow only linearly with cosmic epoch. The instability corresponds to the slow divergence between the variation of the scale factors with cosmic epoch of the model with $\Omega_0 = 1$ and one with slightly greater density. This is the essence of the argument developed by Tolman and Lemaître in the 1930s and developed more generally by Lifshitz in 1946 to the effect that, because the instability develops only algebraically, galaxies could not have formed by gravitational collapse.

The General Solutions

A general solution of (20) for the growth of the density contrast with scale-factor for all pressure-free Friedman world models can be rewritten in terms of the density parameter Ω_0 as follows:

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = \frac{3\Omega_0 H_0^2}{2} a^{-3}\Delta\,, \tag{30}$$

where, in general,

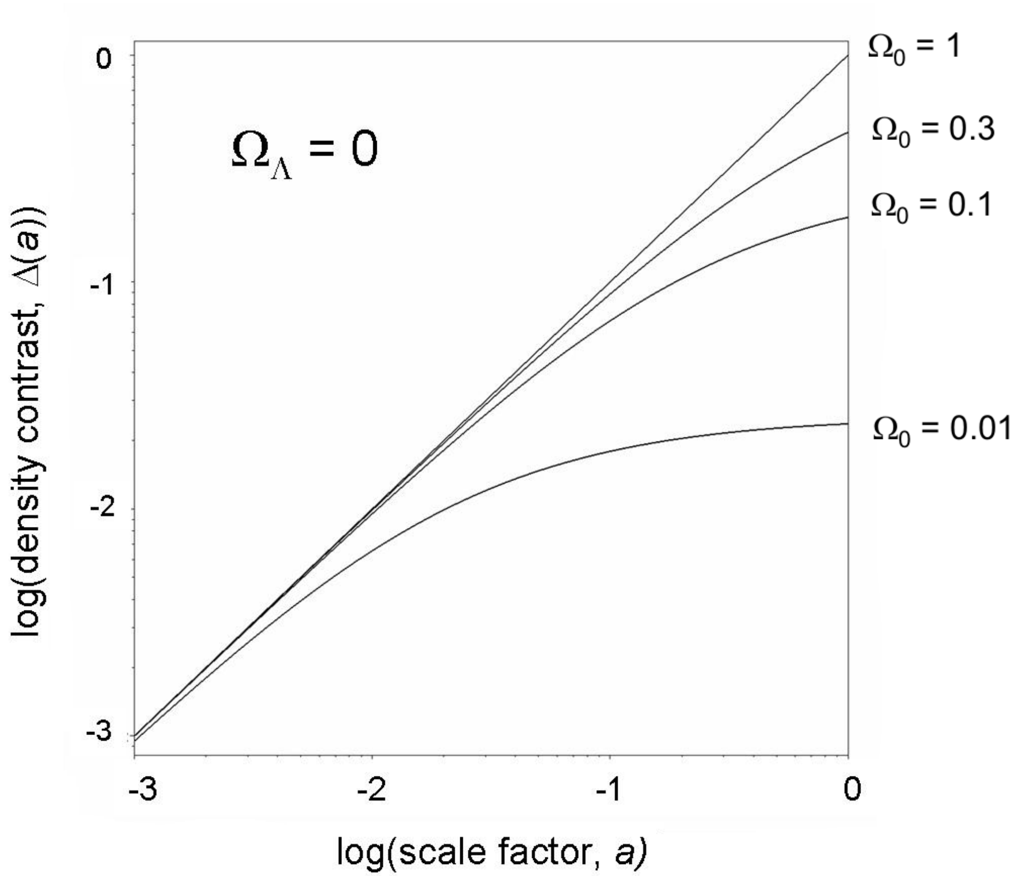
$$\dot{a} = H_0\left[\Omega_0\left(\frac{1}{a} - 1\right) + \Omega_\Lambda(a^2 - 1) + 1\right]^{1/2}\,. \tag{31}$$

The solution for the growing mode can be written as follows:

$$\Delta(a) = \frac{5\Omega_0}{2}\left(\frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t}\right)\int_0^a \frac{\mathrm{d}a'}{(\mathrm{d}a'/\mathrm{d}t)^3}\,, \tag{32}$$

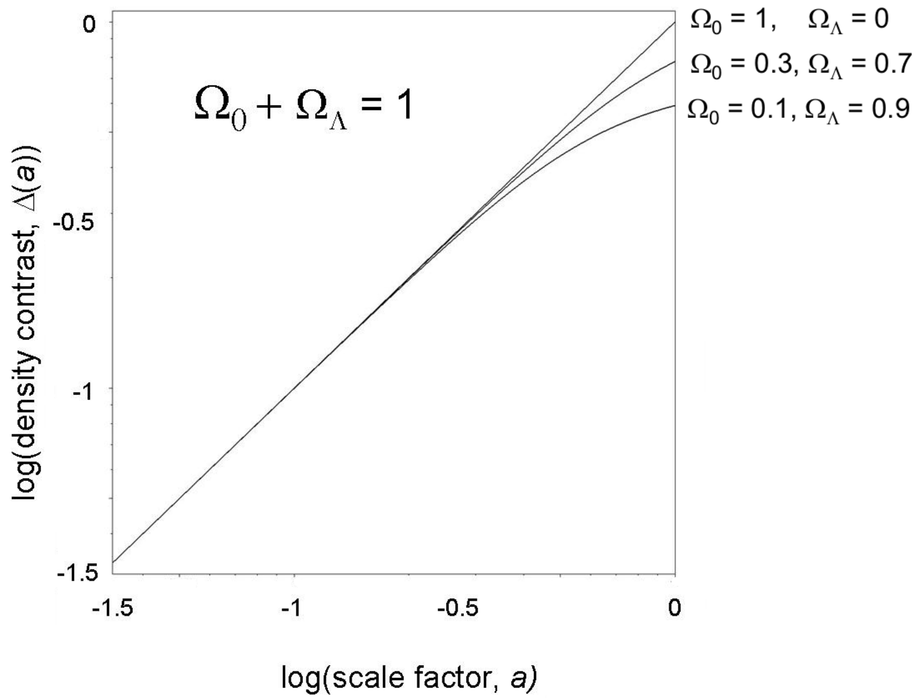
where the constants have been chosen so that the density contrast for the standard critical world model with $\Omega_0 = 1$ and $\Omega_\Lambda = 0$ has unit amplitude at the present epoch, $a = 1$. With this scaling, the density contrasts for all the examples we will consider correspond to $\Delta = 10^{-3}$ at $a = 10^{-3}$. It is simplest to carry out the calculations numerically for a representative sample of world models.

Models with $\Omega_\Lambda = 0$



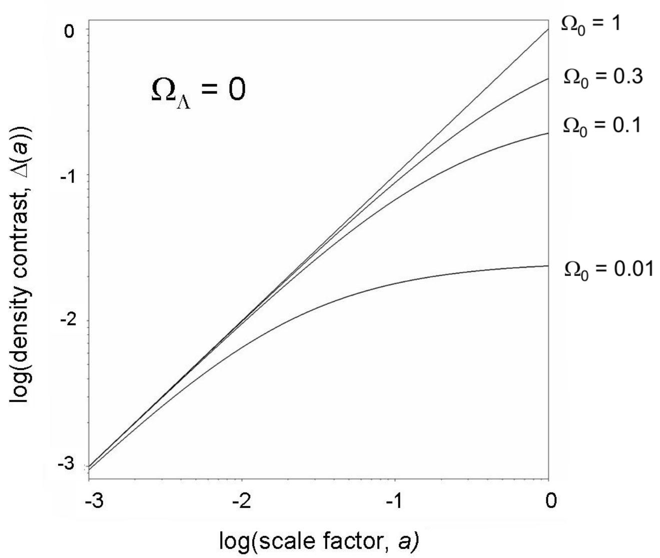
The development of density fluctuations from a scale factor $a = 1/1000$ to $a = 1$ are shown for a range of world models with $\Omega_\Lambda = 0$. These results are consistent with the calculations carried out above, in which it was argued that the amplitudes of the density perturbations vary as $\Delta \propto a$ so long as $\Omega_0 z \gg 1$, but the growth essentially stops at smaller redshifts.

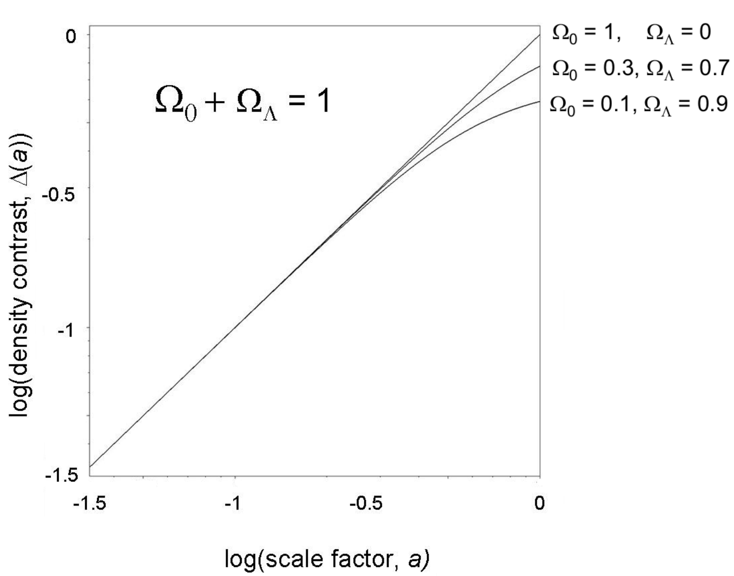
Models with finite Ω_Λ



The models of greatest interest are the flat models for which $(\Omega_0 + \Omega_\Lambda) = 1$, in all cases, the fluctuations having amplitude $\Delta = 10^{-3}$ at $a = 10^{-3}$. The growth of the density contrast is somewhat greater in the cases Ω_0 = 0.1 and 0.3 as compared with the corresponding cases with $\Omega_\Lambda = 0$. The fluctuations continue to grow to greater values of the scale-factor a, corresponding to smaller redshifts, as compared with the models with $\Omega_\Lambda = 0$.

Why are these results so different?





The reason for these differences is that, if $\Omega_\Lambda = 0$, the condition $\Omega_0 z = 1$, corresponds to the change from flat to hyperbolic geometry. This means that neighbouring geodesics are diverging and reduces the strength of the gravitational force.

In the case $\Omega_0 + \Omega_\Lambda = 1$, the geometry is forced to be Euclidean and so the growth continues until the repulsive effect of the Λ term overwhelms the attractive force of gravity. The changeover takes place at much smaller redshifts at $(1+z) \approx \Omega_0^{-1/3}$ if $\Omega_0 \ll 1$.

This is good news if we want to suppress the fluctuations in the Cosmic Microwave Background Radiation.

Peculiar Velocities in the Expanding Universe

The development of velocity perturbations in the expanding Universe can be investigated in the case in which we can neglect pressure gradients so that the velocity perturbations are only driven by the potential gradient $\delta\phi$.

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + 2\left(\frac{\dot{a}}{a}\right)\boldsymbol{u} = -\frac{1}{a^2}\nabla_{\mathrm{c}}\delta\phi \ . \tag{33}$$

In (33), $\boldsymbol{u}$ is the perturbed *comoving* velocity and ∇_{c} is the gradient in comoving coordinates. We split the velocity vectors into components parallel and perpendicular to the gravitational potential gradient, $\boldsymbol{u} = \boldsymbol{u}_{\|} + \boldsymbol{u}_{\perp}$, where $\boldsymbol{u}_{\|}$ is parallel to $\nabla_{\mathrm{c}}\delta\phi$. The velocity associated with $\boldsymbol{u}_{\|}$ is often referred to as *potential motion* since it is driven by the potential gradient. On the other hand, the perpendicular velocity component $\boldsymbol{u}_{\perp}$ is not driven by potential gradients and corresponds to *vortex* or *rotational motions*.

Peculiar Velocities in the Expanding Universe

Rotational Velocities. Consider first the rotational component $\boldsymbol{u}_\perp$. The equation for the peculiar velocity reduces to

$$\frac{\mathrm{d}\boldsymbol{u}_\perp}{\mathrm{d}t} + 2\left(\frac{\dot{a}}{a}\right)\boldsymbol{u}_\perp = 0\ . \tag{34}$$

The solution of this equation is straightforward $\boldsymbol{u}_\perp \propto a^{-2}$. Since $\boldsymbol{u}_\perp$ is a comoving perturbed velocity, the proper velocity is $\delta\boldsymbol{v}_\perp = a\boldsymbol{u}_\perp \propto a^{-1}$. Thus, the rotational velocities decay as the Universe expands.

This is no more than the conservation of angular momentum in an expanding medium, $mvr =$ constant. This poses a grave problem for models of galaxy formation involving primordial turbulence. Rotational turbulent velocities decay and there must be sources of turbulent energy, if the rotational velocities are to be maintained.

Peculiar Velocities in the Expanding Universe

Potential Motions. The development of potential motions can be directly derived from the equation

$$\frac{\mathrm{d}\Delta}{\mathrm{d}t} = -\nabla \cdot \delta \boldsymbol{v} \,, \tag{35}$$

that is, the divergence of the peculiar velocity is proportional to minus the rate of growth of the density contrast. For the case $\Omega_0 = 1$,

$$|\delta v_{\parallel}| = |au| = \frac{H_0 a^{1/2}}{k} \left(\frac{\delta\varrho}{\varrho}\right)_0 = \frac{H_0}{k} \left(\frac{\delta\varrho}{\varrho}\right)_0 (1+z)^{-1/2} \,, \tag{36}$$

where $(\delta\varrho/\varrho)_0$ is the density contrast at the present epoch. Thus, $\delta v_{\parallel} \propto t^{1/3}$.

The peculiar velocities are driven by both the amplitude of the perturbation and its scale. Equation (36) shows that, if $\delta\varrho/\varrho$ is the same on all scales, the peculiar velocities are driven by the smallest values of k, that is, by the perturbations on the largest physical scales. This is an important result for understanding the origin of the peculiar motion of the Galaxy with respect to the frame of reference in which the Microwave Background Radiation is 100% isotropic and of large-scale streaming velocities.

The Relativistic Case

In the radiation-dominated phase of the Big Bang, the primordial perturbations are in a radiation-dominated plasma, for which the relativistic equation of state $p = \frac{1}{3}\varepsilon$ is appropriate.

The equation of energy conservation becomes

$$\frac{\partial \varrho}{\partial t} = -\nabla \cdot \left(\varrho + \frac{p}{c^2}\right) \boldsymbol{v}; \qquad (37)$$

$$\frac{\partial}{\partial t}\left(\varrho + \frac{p}{c^2}\right) = \frac{\dot{p}}{c^2} - \left(\varrho + \frac{p}{c^2}\right)(\nabla \cdot \boldsymbol{v}) . \qquad (38)$$

Substituting $p = \frac{1}{3}\varrho c^2$ into (37) and (38), the relativistic continuity equation is obtained:

$$\frac{\mathrm{d}\varrho}{\mathrm{d}t} = -\frac{4}{3}\varrho(\nabla \cdot \mathrm{v}) . \qquad (39)$$

Euler's equation for the acceleration of an element of the fluid in the gravitational potential ϕ becomes

$$\left(\varrho + \frac{p}{c^2}\right)\left[\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}\right] = -\nabla p - \left(\varrho + \frac{p}{c^2}\right)\nabla\phi . \qquad (40)$$

The Relativistic Case

If we neglect the pressure gradient term, (40) reduces to the familiar equation

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\nabla\phi \, . \tag{41}$$

Finally, the differential equation for the gravitational potential ϕ becomes

$$\nabla^2\phi = 4\pi G\left(\varrho + \frac{3p}{c^2}\right) \, . \tag{42}$$

For a fully relativistic gas, $p = \frac{1}{3}\varrho c^2$ and so

$$\nabla^2\phi = 8\pi G\varrho \, . \tag{43}$$

The net result is that the equations for the evolution of the perturbations in a relativistic gas are of similar mathematical form to the non-relativistic case. The same type of analysis which was carried out above leads to the following equation

$$\boxed{\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} = \Delta\left(\frac{32\pi G\varrho}{3} - k^2 c_\mathrm{s}^2\right) \, .} \tag{44}$$

The Relativistic Case

The relativistic expression for the Jeans' length is found by setting the right-hand side equal to zero,

$$\lambda_{\mathrm{J}} = \frac{2\pi}{k_{\mathrm{J}}} = c_{\mathrm{s}} \left(\frac{3\pi}{8G\varrho} \right)^{1/2} , \tag{45}$$

where $c_{\mathrm{s}} = c/\sqrt{3}$ is the relativistic sound speed. The result is similar to the standard expression for the Jeans' length.

Neglecting the pressure gradient terms in (44), the following differential equation for the growth of the instability is obtained

$$\frac{\mathrm{d}^2\Delta}{\mathrm{d}t^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}t} - \frac{32\pi G\varrho}{3}\Delta = 0 \, . \tag{46}$$

We again seek solutions of the form $\Delta = at^n$, recalling that in the radiation-dominated phases, the scale factor-cosmic time relation is given by $a \propto t^{1/2}$. We find solutions $n = \pm 1$. Hence, for wavelengths $\lambda \gg \lambda_{\mathrm{J}}$, the growing solution corresponds to

$$\Delta \propto t \propto a^2 \propto (1+z)^{-2} \, . \tag{47}$$

Thus, once again, the unstable mode grows algebraically with cosmic time.

The Basic Problem of Structure Formation

Let us summarise the implications of the key results derived above. Throughout the matter- dominated era, the growth rate of perturbations on physical scales much greater than the Jeans' length is

$$\Delta = \frac{\delta\varrho}{\varrho} \propto a = \frac{1}{1+z} \, . \qquad (48)$$

Since galaxies and astronomers certainly exist at the present day $z = 0$, it follows that $\Delta \geq 1$ at $z = 0$ and so, at the last scattering surface, $z \sim 1,000$, fluctuations must have been present with amplitude at least $\Delta = \delta\varrho/\varrho \geq 10^{-3}$.

- The slow growth of density perturbations is the source of a fundamental problem in understanding the origin of galaxies – large-scale structures did not condense out of the primordial plasma by exponential growth of infinitesimal statistical perturbations.
- Because of the slow development of the density perturbations, we have the opportunity of studying the formation of structure on the last scattering surface at a redshift $z \sim 1,000$.

Matter and Radiation in the Universe

The Cosmic Microwave Background Radiation provides by far the greatest contribution to the energy density of radiation in intergalactic space. Comparing the inertial mass density in the radiation and the matter, we find

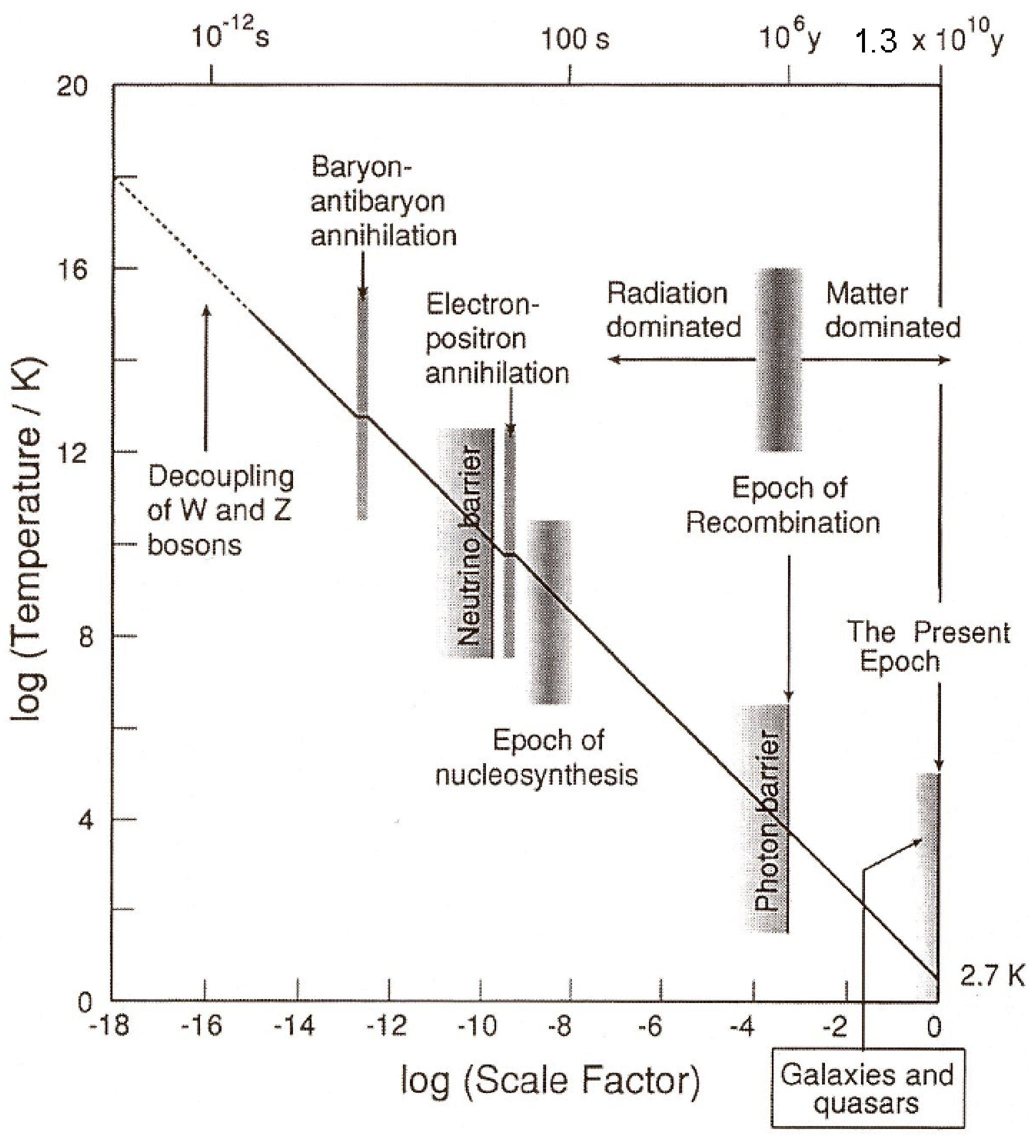
$$\frac{\varrho_{\rm r}}{\varrho_{\rm m}} = \frac{aT^4(z)}{\Omega_0 \varrho_{\rm c}(1+z)^3 c^2} = \frac{2.48 \times 10^{-5}(1+z)}{\Omega_0 h^2}\,. \tag{49}$$

Thus, at redshifts $z \geq 4 \times 10^4 \Omega_0 h^2$, the Universe was certainly *radiation-dominated*, even before we take account of the contribution of the three types of neutrino to the inertial mass density during the radiation-dominated phase, and the dynamics are described by the relation, $a \propto t^{1/2}$. According to this analysis, the Universe is *matter-dominated* at redshifts $z \leq 4 \times 10^4 \Omega_0 h^2$ and the dynamics are described by the standard Friedman models, $a \propto t^{2/3}$ provided $\Omega_0 z \gg 1$.

The present *photon-to-baryon ratio* is another key cosmological parameter. Assuming $T = 2.728$ K,

$$\frac{N_\gamma}{N_{\rm B}} = \frac{3.6 \times 10^7}{\Omega_{\rm B} h^2}\,. \tag{50}$$

Summary of the Thermal History of the Universe



This diagram summarises the key epochs in the thermal history of the Universe. The key epochs are

- The epoch of recombination.
- The epoch of equality of matter and radiation.

The Epoch of Recombination

At a redshift $z \approx 1500$, the radiation temperature of the Cosmic Microwave Background Radiation was $T \approx 4,000$ K and then there were sufficient photons with energies $E = h\nu \geq 13.6$ eV in the tail of the Planck distribution to ionise most of the neutral hydrogen present in the intergalactic medium.

It is a useful calculation to work out the fraction of photons in the high frequency tail of the Planck distribution, that is, in the *Wien region* of the spectrum, with energies $h\nu \geq E$ in the limit $h\nu \gg kT$.

$$n(\geq E) = \int_{E/h}^{\infty} \frac{8\pi\nu^2}{c^3} \frac{\mathrm{d}\nu}{\mathrm{e}^{h\nu/kT}} = \frac{1}{\pi^2}\left(\frac{2\pi kT}{hc}\right)^3 \mathrm{e}^{-x}(x^2 + 2x + 2)\,, \qquad (51)$$

where $x = h\nu/kT$. Now, the total number density of photons in a black-body spectrum at temperature T is

$$N = 0.244\left(\frac{2\pi kT}{hc}\right)^3 \mathrm{m}^{-3}\,. \qquad (52)$$

The Epoch of Recombination

Therefore, the fraction of the photons of the black-body spectrum with energies greater than E is

$$\frac{n(\geq E)}{n_{\text{ph}}} = \frac{\text{e}^{-x}(x^2 + 2x + 2)}{0.244\pi^2} . \tag{53}$$

Roughly speaking, the intergalactic gas will be ionised, provided there are as many ionising photons with $h\nu \geq 13.6$ eV as there are hydrogen atoms, that is, we need only one photon in $3.6 \times 10^7/\Omega_0 h^2$ of the photons of the Cosmic Microwave Background Radiation to have energy greater than 13.6 eV to ionise the gas. For illustrative purposes, let us take the ratio to be one part in 10^9 Then, we need to solve

$$\frac{1}{10^9} = \frac{\text{e}^{-x}(x^2 + 2x + 2)}{0.244\pi^2} . \tag{54}$$

We find $x = E/kT \approx 26.5$. There are so many photons relative to hydrogen atoms that the temperature of the radiation can be 26.5 times less than that found from setting $E = kT$ and there are still sufficient photons with energy $E \geq 13.6$ eV to ionise the gas. Therefore, the intergalactic gas was largely ionised at a temperature of 150,000/26.5 K $\approx 5,600$ K.

The Ionisation of the Intergalactic Gas Through the Epoch of Recombination

The optical depth of the intergalactic gas increases rapidly with redshift once the gas becomes fully ionised. Temperature fluctuations which originate at redshifts greater than the redshift of recombination are damped out by Thomson scattering. The fluctuations we observe originate in a rather narrow redshift range about that at which the optical depth of the intergalactic gas is unity.

At the epoch of recombination, the plasma was 50% ionised when the temperature of the background radiation was about 4,000 K. Photons emitted in the recombination of hydrogen atoms must have energies $h\nu \geq h\nu_\alpha$, where ν_α is the frequency of the Lyman-α transition which has wavelength 121.6 nm. These photons can either reionise other hydrogen atoms directly, or else raise them to an excited state H^*, from which the electron can be ejected by the much more plentiful soft photons in the black-body spectrum.

The Lyman-α photons are destroyed by the *two-photon process* in which two photons are liberated from the 2s state of hydrogen in a rare quadrupole transition with spontaneous transition probability $w = 8.23$ sec^{-1}.

The Probability Distribution of Last Scattering

Jones and Wyse (1985) provided a convenient analytic expression for the degree of ionisation through the critical redshift range:

$$x = 2.4 \times 10^{-3} \frac{(\Omega_0 h^2)^{1/2}}{\Omega_B h^2} \left(\frac{z}{1000}\right)^{12.75} . \tag{55}$$

Ω_0 is the density parameter for the Universe as a whole and Ω_B the density parameter of baryons. The optical depth of the intergalactic gas at redshifts $z \sim 1000$ is

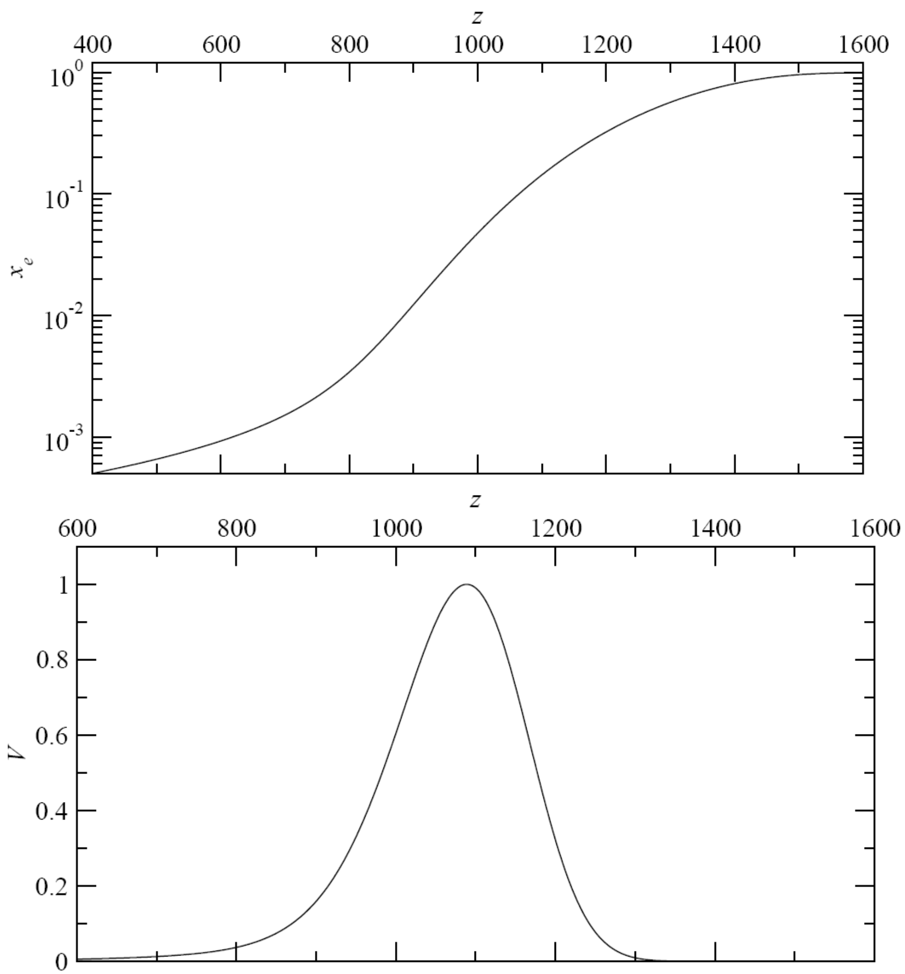
$$\tau = 0.37 \left(\frac{z}{1000}\right)^{14.25} . \tag{56}$$

Because of the enormously strong dependence upon redshift, the optical depth of the intergalactic gas is always unity very close to a redshift of 1070, independent of the exact values of Ω_0, Ω_B and h. This probability distribution for the range of redshifts from which the photons of the background radiation we observe today were last scattered is given by

$$p(z)\,\mathrm{d}z = e^{-\tau} \frac{\mathrm{d}\tau}{\mathrm{d}z}\,\mathrm{d}z , \tag{57}$$

which can be closely approximated by a Gaussian distribution with mean redshift 1070 and standard deviation $\sigma = 80$ in redshift.

The Probability Distribution of Last Scattering



Improved calculations were carried out by Chluba and Sunyaev (2006) in which the populations of the excited states departed from their equilibrium distributions.

The top diagram shows the fractional ionisation and the lower diagram the probability distribution of last scattering.

Fluctuations at the Last Scattering Layer

The physical scale at the present epoch corresponding to the thickness of the last scattering layer is given by

$$\mathrm{d}r = \frac{c}{H_0} \frac{\mathrm{d}z}{z^{3/2} \Omega_0^{1/2}} \, . \tag{58}$$

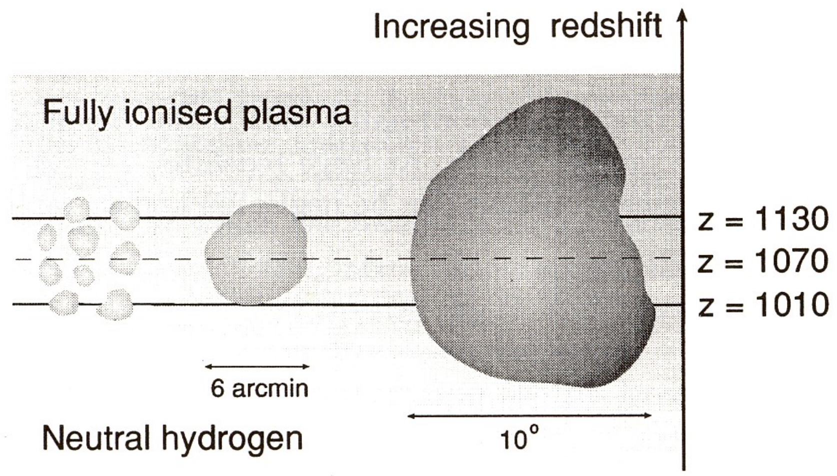
If we take the thickness of the last scattering layer to correspond to a redshift interval $\Delta z = 195$ at $z = 1090$, this is equivalent to a physical scale of $16.2(\Omega_0 h^2)^{-1/2} = 42$ Mpc at the present epoch. The mass contained within this scale is $M \approx 6 \times 10^{14} (\Omega_0 h^2)^{-1/2} \, M_\odot \sim 1.6 \times 10^{15} \, M_\odot$, corresponding roughly to the mass of a cluster of galaxies.

The comoving scale of $d = 16.2(\Omega_0 h^2)^{-1/2}$ Mpc corresponds to a proper distance $d/(1+z)$ at redshift z and hence to an angular size

$$\theta = \frac{d(1+z)}{D} = \frac{16.2(\Omega_0 h^2)^{-1/2}}{D_{\mathrm{Mpc}}} = 5.8 \Omega_0^{1/2} \, \mathrm{arcmin} = 3.2 \, \mathrm{arcmin} \, , \tag{59}$$

since $D = 2c/H_0 \Omega_0$, if $\Omega_0 z \gg 1$.

Perturbations on the Last Scattering Layer



The diagram shows schematically the size of various small perturbations compared with the thickness of the last scattering layer. On very large scales, the perturbations are very much larger than the thickness of the layer. On scales less than clusters of galaxies, many perturbations overlap, reducing the amplitude of the perturbations.

The Sound Speed as a Function of Cosmic Epoch

All sound speeds are proportional to the square root of the ratio of the pressure which provides the restoring force to the inertial mass density of the medium. The speed of sound c_{s} is given by

$$c_{\mathrm{s}}^2 = \left(\frac{\partial p}{\partial \varrho}\right)_S , \qquad (60)$$

where the subscript S means 'at constant entropy', that is, we consider adiabatic sound waves. From the epoch when the energy densities of matter and radiation were equal to beyond the epoch of recombination, the dominant contributors to p and ϱ change dramatically as the Universe changes from being radiation- to matter-dominated. The sound speed can then be written

$$c_{\mathrm{s}}^2 = \frac{(\partial p/\partial T)_{\mathrm{r}}}{(\partial \varrho/\partial T)_{\mathrm{r}} + (\partial \varrho/\partial T)_{\mathrm{m}}}, \qquad (61)$$

where the partial derivatives are taken at constant entropy. It is straightforward to show that this reduces to the following expression:

$$c_{\mathrm{s}}^2 = \frac{c^2}{3}\frac{4\varrho_{\mathrm{r}}}{4\varrho_{\mathrm{r}} + 3\varrho_{\mathrm{m}}} . \qquad (62)$$

The Sound Speed as a Function of Cosmic Epoch

Thus, in the radiation-dominated era, $z \gg 4 \times 10^4 \Omega_0 h^2$, $\varrho_\mathrm{r} \gg \varrho_\mathrm{m}$ and the speed of sound tends to the relativistic sound speed, $c_\mathrm{s} = c/\sqrt{3}$.

At smaller redshifts, the sound speed decreases as the contribution of the inertial mass density of the matter becomes more important. Between the epoch of equality of the matter and radiation energy densities and the epoch of the recombination, the pressure of sound waves is provided by the radiation, but the inertia is provided by the matter. Thus, the speed of sound decreases to

$$c_\mathrm{s} = \left(\frac{4c^2}{9}\frac{\varrho_\mathrm{r}}{\varrho_\mathrm{m}}\right)^{1/2} = \left[\frac{4aT_0^4(1+z)}{9\Omega_\mathrm{m}\varrho_\mathrm{c}}\right]^{1/2} = \frac{10^6 z^{1/2}}{(\Omega_\mathrm{m} h^2)^{1/2}}\ \mathrm{m\ s^{-1}}\,. \qquad (63)$$

After recombination, the sound speed is the thermal sound speed of the matter which, because of the close coupling between the matter and the radiation, has temperature $T_\mathrm{r} = T_\mathrm{m}$ at redshifts $z \geq 550\, h^{2/5}\Omega_0^{1/5}$. Thus, at a redshift of 500, the temperature of the gas was 1300 K.

The Damping of Sound Waves

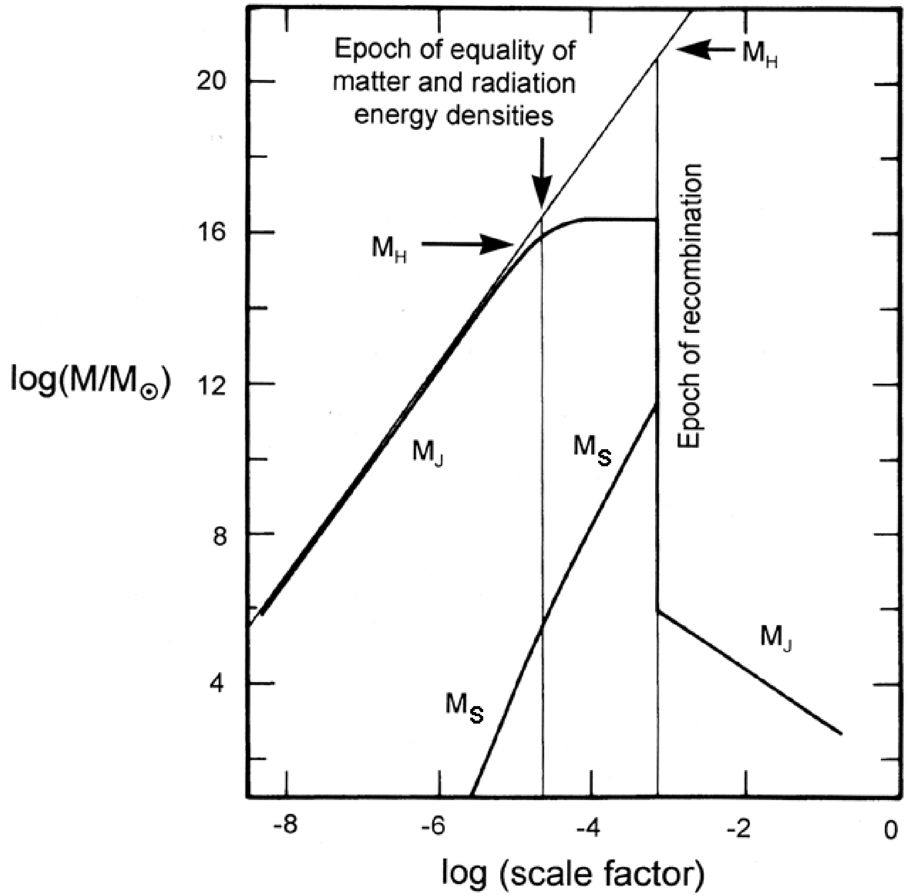
Although the matter and radiation are closely coupled throughout the pre-recombination era, the coupling is not perfect and radiation can diffuse out of the density perturbations. Since the radiation provides the restoring force for support for the perturbation, the perturbation is damped out if the radiation has time to diffuse out of it. This process is often referred to as *Silk damping*.

At any epoch, the mean free path for scattering of photons by electrons is $\lambda = (N_\mathrm{e}\sigma_\mathrm{T})^{-1}$, where $\sigma_\mathrm{T} = 6.665 \times 10^{-29}$ m^2 is the Thomson cross-section. The distance which the photons can diffuse is

$$r_\mathrm{D} \approx (Dt)^{1/2} = \left(\tfrac{1}{3}\lambda c t\right)^{1/2} , \qquad (64)$$

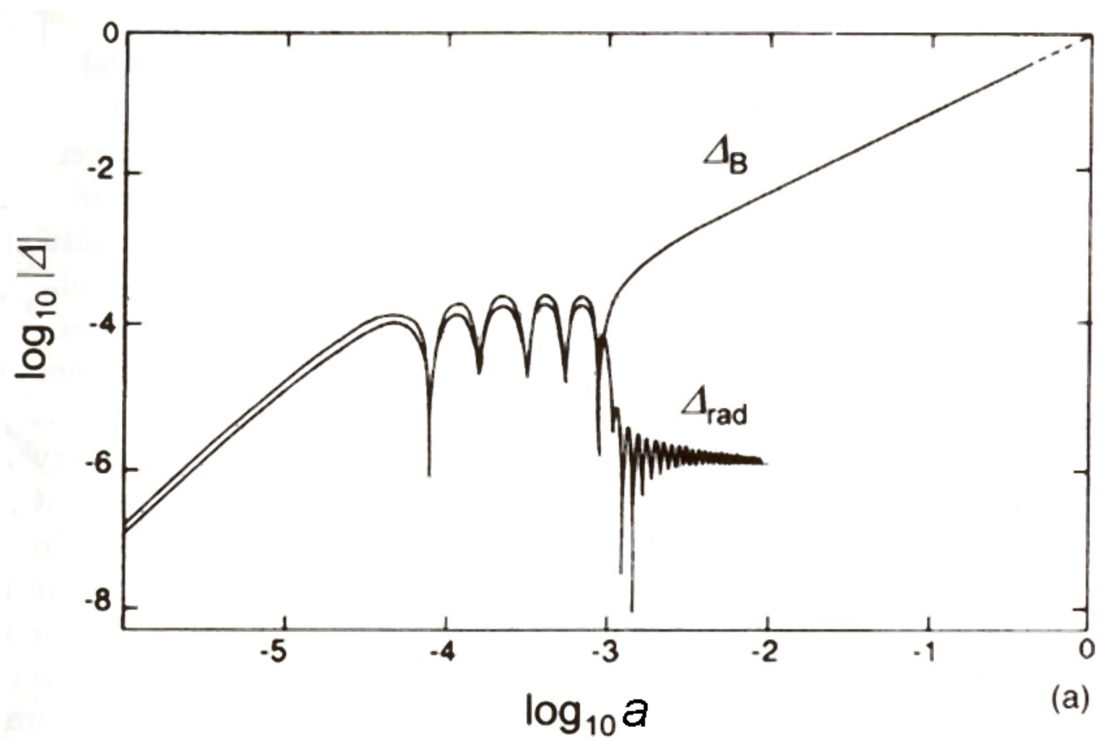
where t is cosmic time. The baryonic mass within this radius, $M_\mathrm{D} = (4\pi/3) r_\mathrm{D}^3 \varrho_\mathrm{B}$, can now be evaluated for the pre-recombination era.

The Simple Baryonic Picture



We can put together all these ideas to develop the simplest picture of galaxy formation. This is the simplest baryonic picture. It includes many of the features which will reappear in the ΛCDM picture. The diagram shows how the horizon mass M_{H}, the Jeans mass M_{J} and the Silk Mass M_{D} change with scale factor a.

The Simple Baryonic Picture



This diagram, from Coles and Lucchin (1995), shows schematically how structure develops in a purely baryonic Universe. The problem is that the temperature fluctuations on the last scattering surface are expected to be at least $\Delta T/T \sim 10^{-3}$, far in excess of the observed limits.

The solution to this problem came with the realisation that the dark matter is the dominant contribution to Ω_0.

Instabilities in the Presence of Dark Matter

Neglecting the internal pressure of the fluctuations, the expressions for the density contrasts in the baryons and the dark matter, Δ_B and Δ_D respectively, can be written as a pair of coupled equations

$$\ddot{\Delta}_B + 2\left(\frac{\dot{a}}{a}\right)\dot{\Delta}_B = A\varrho_B\Delta_B + A\varrho_D\Delta_D \, , \tag{65}$$

$$\ddot{\Delta}_D + 2\left(\frac{\dot{a}}{a}\right)\dot{\Delta}_D = A\varrho_B\Delta_B + A\varrho_D\Delta_D \, . \tag{66}$$

Let us find the solution for the case in which the dark matter has $\Omega_0 = 1$ and the baryon density is negligible compared with that of the dark matter. Then (65) reduces to the equation for which we have already found the solution $\Delta_D = Ba$ where B is a constant. Therefore, the equation for the evolution of the baryon perturbations becomes

$$\ddot{\Delta}_B + 2\left(\frac{\dot{a}}{a}\right)\dot{\Delta}_B = 4\pi G\varrho_D Ba \, . \tag{67}$$

Instabilities in the Presence of Dark Matter

Since the background model is the critical model, equation (67) simplifies to

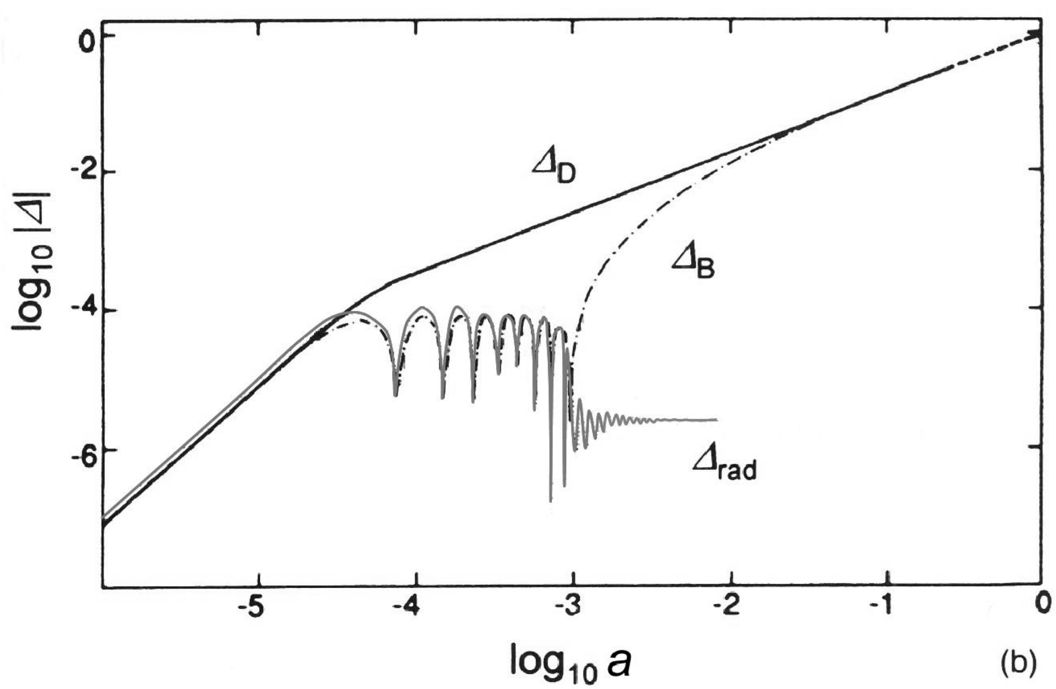
$$a^{3/2}\frac{\mathrm{d}}{\mathrm{d}a}\left(a^{-1/2}\frac{\mathrm{d}\Delta}{\mathrm{d}a}\right)+2\frac{\mathrm{d}\Delta}{\mathrm{d}a}=\tfrac{3}{2}B\,. \tag{68}$$

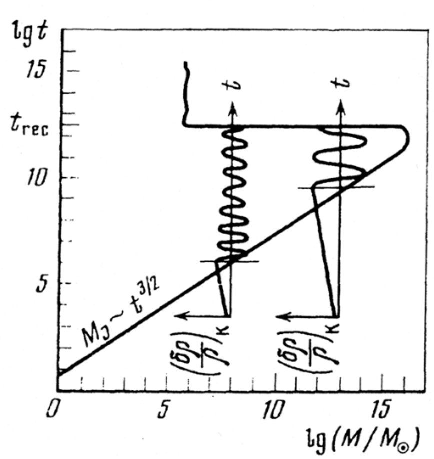
The solution, $\Delta = B(a-a_0)$, satisfies (68). This result has the following significance. Suppose that, at some redshift z_0, the amplitude of the baryon fluctuations is very small, that is, very much less than that of the perturbations in the dark matter. The above result shows how the amplitude of the baryon perturbation develops subsequently under the influence of the dark matter perturbations. In terms of redshift we can write

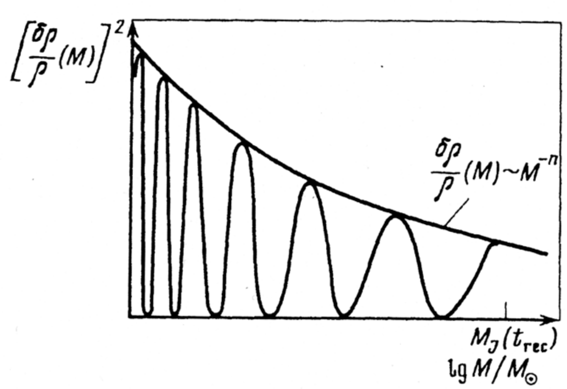
$$\Delta_{\mathrm{B}} = \Delta_{\mathrm{D}}\left(1-\frac{z}{z_0}\right)\,. \tag{69}$$

Thus, the amplitude of the perturbations in the baryons grows rapidly to the same amplitude as that of the dark matter perturbations. The baryons fall into the dark matter perturbations and rapidly attain amplitudes the same as those of the dark matter.

The Cold Dark Matter Picture







This diagram shows how structure develops in a cold dark matter dominated Universe. The amplitudes of the baryonic perturbations were very much smaller than those in the cold dark matter at the epoch of recombination.

Note also the origin of the acoustic peaks in the predicted mass spectrum (from Sunyaev and Zeldovich 1970).

This is the favoured model for the formation structure.

The Initial Power-Spectrum

The smoothness of the two-point correlation function for galaxies suggest that the spectrum of initial fluctuations must have been very broad with no preferred scales and it is therefore natural to begin with a power spectrum of power-law form

$$P(k) = |\Delta_k|^2 \propto k^n \,. \tag{70}$$

The correlation function $\xi(r)$ should then have the form

$$\xi(r) \propto \int \frac{\sin kr}{kr} k^{(n+2)} \, \mathrm{d}k \,. \tag{71}$$

Because the function $\sin kr/kr$ has value unity for $kr \ll 1$ and decreases rapidly to zero when $kr \gg 1$, we can integrate k from 0 to $k_{\max} \approx 1/r$ to estimate the dependence of the amplitude of the correlation function on the scale r.

$$\xi(r) \propto r^{-(n+3)} \,. \tag{72}$$

Since the mass of the fluctuation is proportional to r^3, this result can also be written in terms of the mass within the fluctuations on the scale r, $M \sim \varrho r^3$.

$$\xi(M) \propto M^{-(n+3)/3} \,. \tag{73}$$

The Initial Power-Spectrum

Finally, to relate ξ to the root-mean-square density fluctuation on the mass scale M, $\Delta(M)$, we take the square root of ξ, that is,

$$\Delta(M) = \frac{\delta\varrho}{\varrho}(M) = \langle\Delta^2\rangle^{1/2} \propto M^{-(n+3)/6} \,. \qquad (74)$$

This spectrum has the important property that the density contrast $\Delta(M)$ had the same amplitude on all scales when the perturbations came through their particle horizons, provided $n = 1$. Let us illustrate how this comes about.

Before the perturbations came through their particle horizons and before the epoch of equality of matter and radiation energy densities, the density perturbations grew as $\Delta(M) \propto a^2$, although the perturbation to the gravitational potential was frozen-in. Therefore, the development of the spectrum of density perturbations can be written

$$\Delta(M) \propto a^2 \, M^{-(n+3)/6} \,. \qquad (75)$$

The Initial Power-Spectrum

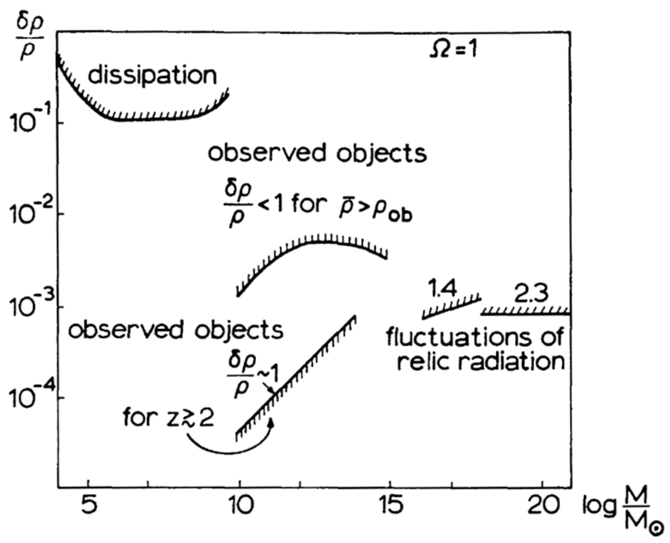
A perturbation of scale r came through the horizon when $r \approx ct$, and so the mass of dark matter within it was $M_{\mathrm{D}} \approx \varrho_{\mathrm{D}}(ct)^3$. During the radiation dominated phases, $a \propto t^{1/2}$ and the number density of dark matter particles, which will eventually form bound structures at $z \sim 0$, varied as $N_{\mathrm{D}} \propto a^{-3}$.

Therefore, the horizon dark matter mass increased as $M_{\mathrm{H}} \propto a^3$, or, $a \propto M_{\mathrm{H}}^{1/3}$. The mass spectrum $\Delta(M)_{\mathrm{H}}$ when the fluctuations came through the horizon at different cosmic epochs was

$$\Delta(M)_{\mathrm{H}} \propto M^{2/3}\, M^{-(n+3)/6} = M^{-(n-1)/6}\,. \tag{76}$$

Thus, if $n = 1$, the density perturbations $\Delta(M) = \delta\varrho/\varrho(M)$ all had the same amplitude when they came though their particle horizons during the radiation-dominated era.

The Harrison–Zeldovich Power Spectrum



The constraints on the form of the perturbation spectrum in 1971 derived by Sunyaev and Zeldovich.

They put in what was needed to produce the observed structures today.

Sunyaev and Zeldovich used a variety of constraints to derive the form of the initial power-spectrum of density perturbations as they came through the horizon. They found a scale-invariant spectrum $\delta\varrho/\varrho = 10^{-4}$ on mass scales from 10^5 to $10^{20}\ M_\odot$.

Harrison studied the form the primordial spectrum must have in order to prevent the overproduction of excessively large amplitude perturbations on small and large scales. A power spectrum of the form

$$P(k) \propto k \qquad (77)$$

does not diverge on large physical scales and so is consistent with the observed large-scale isotropy of the Universe.

Processing of the Initial Power Spectrum

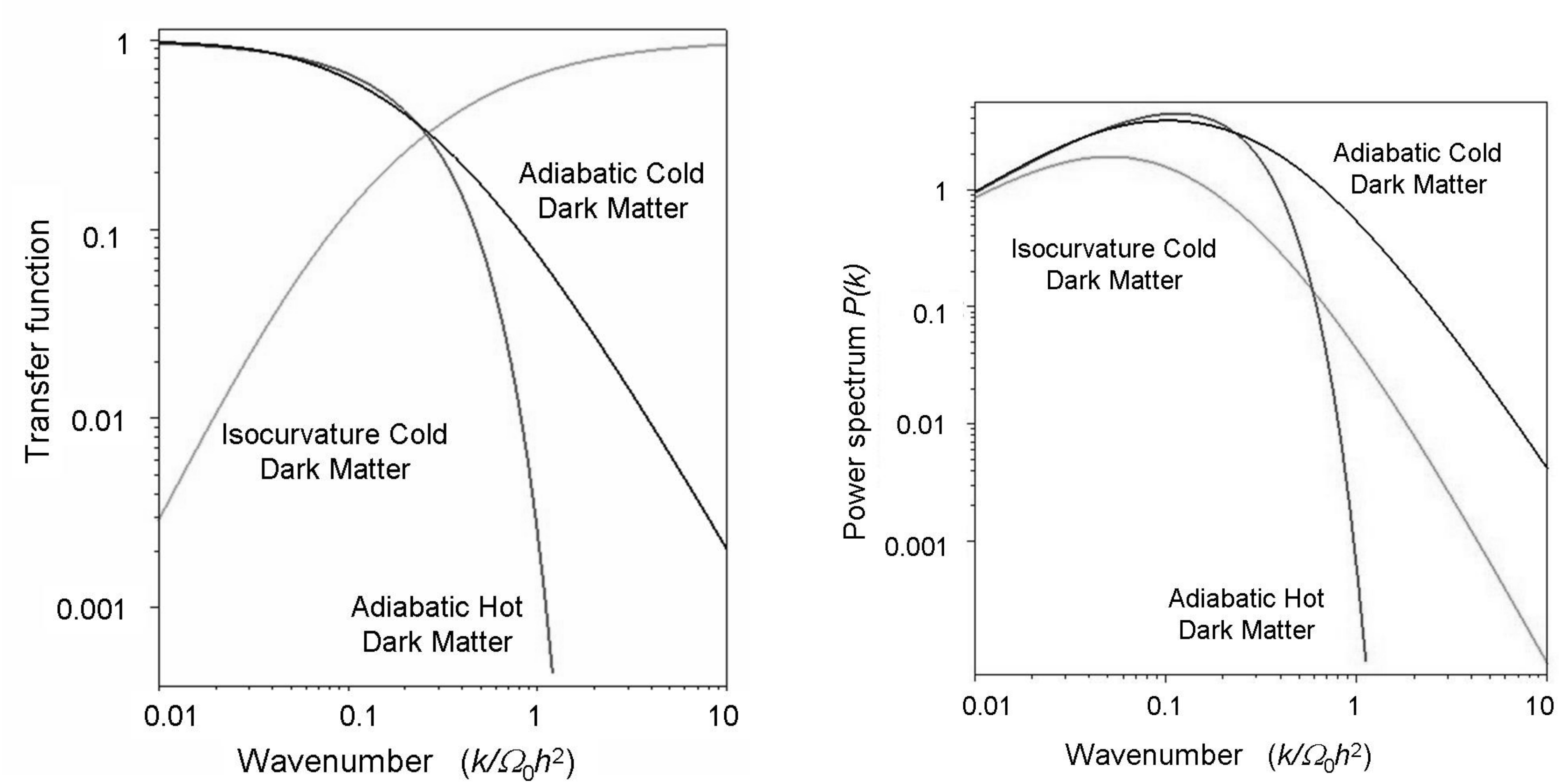
We do not observe the initial power-spectrum except on the largest physical scales. The *transfer function* $T(k)$ which describes how the shape of the initial power-spectrum $\Delta_k(z)$ in the dark matter is modified by different physical processes through the relation

$$\Delta_k(z = 0) = T(k)\, f(z)\, \Delta_k(z)\ . \tag{78}$$

$\Delta_k(z = 0)$ is the power spectrum at the present epoch and $f(z) \propto a \propto t^{2/3}$ is the linear growth factor between the scale factor at redshift z and the present epoch in the matter dominated era.

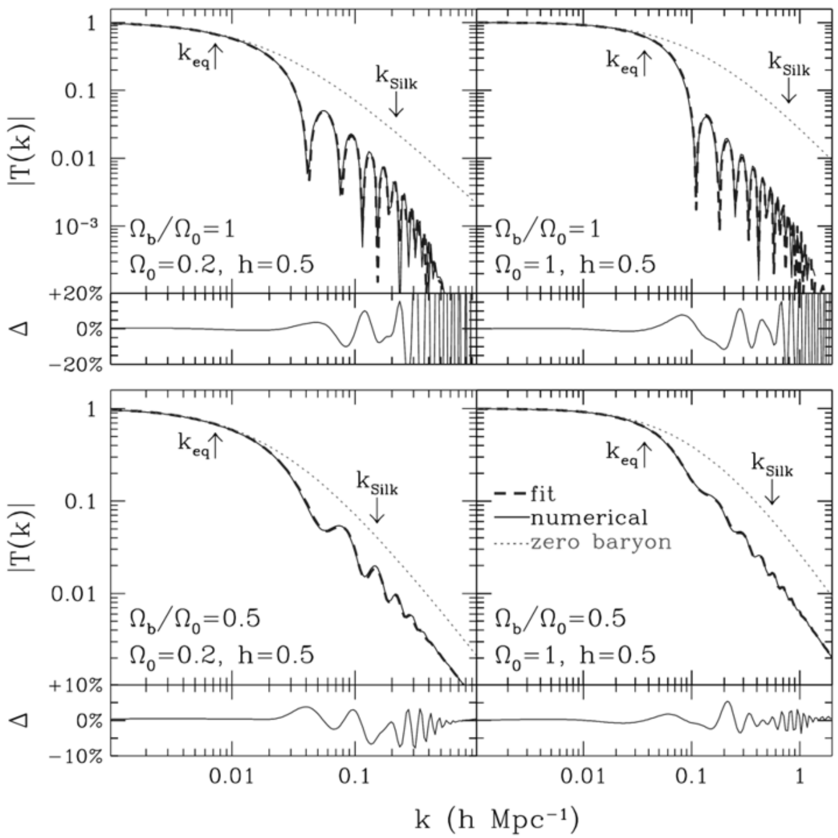
The form of the transfer function is largely determined by the fact that there is a delay in the growth of the perturbations between the time when they came through the horizon and began to grow again. In the standard cold dark matter picture, this is associated with the fact that before the epoch of equality of matter and radiation, the oscillations in the photon-baryon plasma were dynamically more important than those in the dark matter.

The Processed Harrison–Zeldovich Power Spectrum



Notice that on very large scales (small wavenumbers) the spectrum is unprocessed. On the scale of galaxies and clusters, the spectrum has been strongly modified.

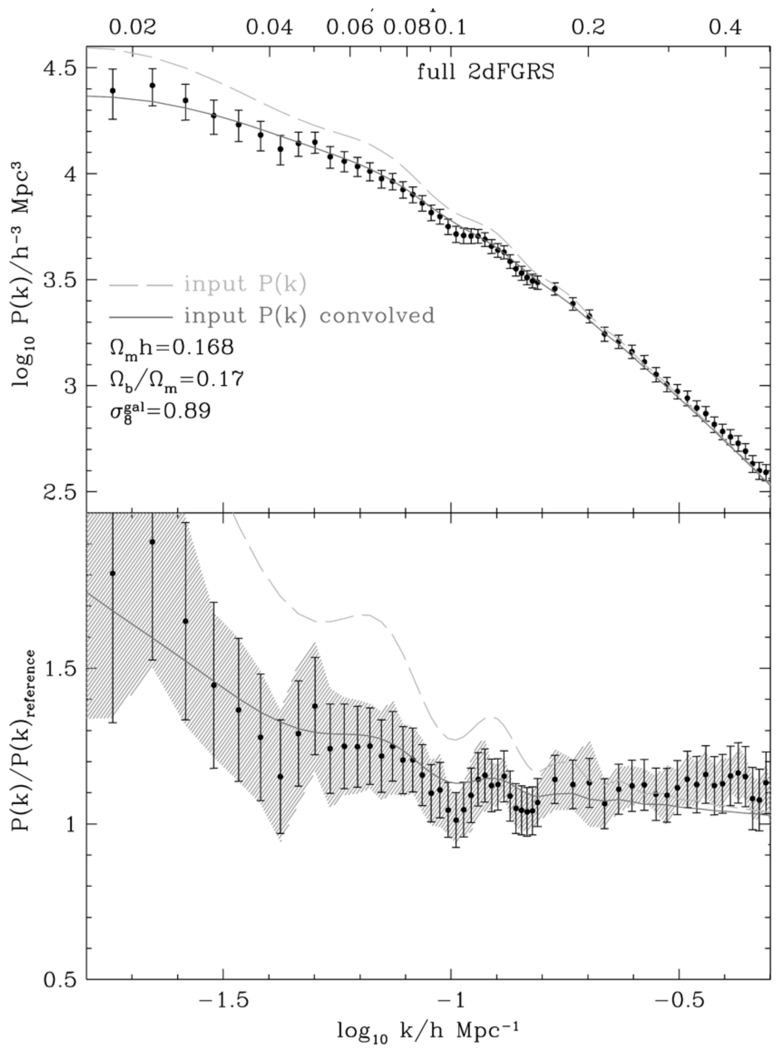
The Processed Harrison–Zeldovich Power Spectrum
Adding in the Baryons



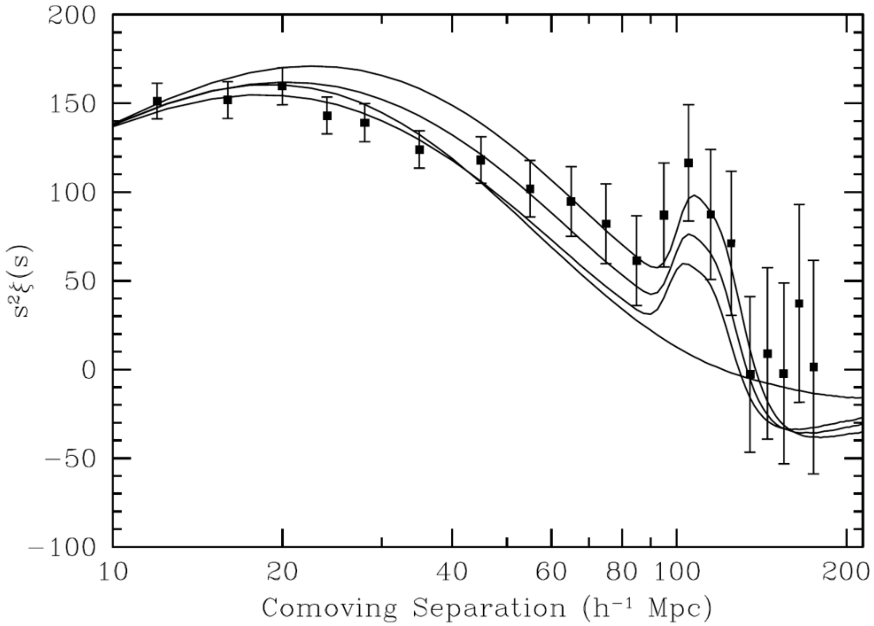
Four examples of the transfer functions for models of structure formation with baryons only (top pair of diagrams) and with mixed cold and baryonic models (bottom pair of diagrams) by Eisenstein and Hu. The numerical results are shown as solid lines and their fitting functions by dashed lines. The lower small boxes in each diagram show the percentage residuals to their fitting functions, which are always less than 10%.

The Acoustic Oscillations in the Galaxy Distribution

AAT 2dF galaxy survey



SDSS galaxy survey



The Non-linear Collapse of Density Perturbations

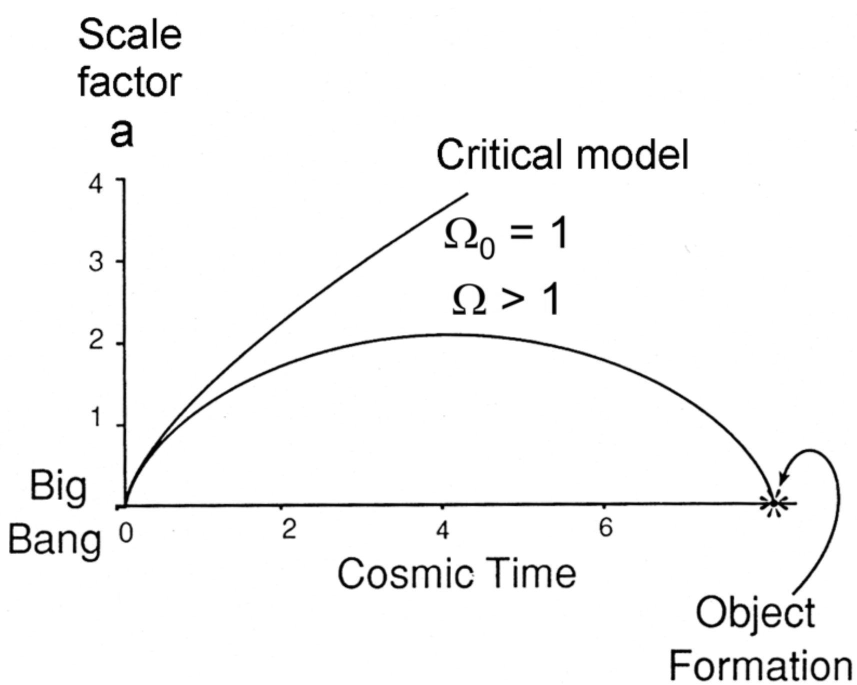
The collapse of a uniform spherical density perturbation in an otherwise uniform Universe can be worked out exactly, a model sometimes referred to as *spherical top-hat collapse*. The dynamics are the same as those of a closed Universe with $\Omega_0 > 1$. The variation of the scale factor of the perturbation a_p is given by the parametric solution

$$\begin{aligned} a_\mathrm{p} &= A(1 - \cos\theta) \qquad t = B(\theta - \sin\theta)\ , \\ A &= \frac{\Omega_0}{2(\Omega_0 - 1)} \quad \text{and} \quad B = \frac{\Omega_0}{2H_0(\Omega_0 - 1)^{3/2}}\ . \end{aligned}$$

The perturbation reached maximum radius at $\theta = \pi$ and then collapsed to infinite density at $\theta = 2\pi$. The perturbation stopped expanding, $\dot{a}_\mathrm{p} = 0$, and separated out of the expanding background at $\theta = \pi$. This occurred when the scale factor of the perturbation was $a = a_\mathrm{max}$, where

$$a_\mathrm{max} = 2A = \frac{\Omega_0}{\Omega_0 - 1} \quad \text{at time} \quad t_\mathrm{max} = \pi B = \frac{\pi\Omega_0}{2H_0(\Omega_0 - 1)^{3/2}}\ . \qquad (79)$$

Perturbing the Friedman solutions



This result indicates why density perturbations grow only linearly with cosmic epoch. The instability corresponds to the slow divergence between the variation of the scale factors with cosmic epoch of the model with $\Omega_0 = 1$ and one with slightly greater density. This is the essence of the argument developed by Tolman and Lemaître in the 1930s and developed more generally by Lifshitz in 1946 to the effect that, because the instability develops only algebraically, galaxies could not have formed by gravitational collapse.

The Non-linear Collapse of Density Perturbations

The density of the perturbation at maximum scale factor ρ_{max} can now be related to that of the background ρ_0, which, for illustrative purposes, we take to be the critical model, $\Omega_0 = 1$. Recalling that the density within the perturbation was Ω_0 times that of the background model to begin with,

$$\frac{\rho_{\mathrm{max}}}{\rho_0} = \Omega_0 \left(\frac{a}{a_{\mathrm{max}}}\right)^3 = 9\pi^2/16 = 5.55 \,, \qquad (80)$$

where the scale factor of the background model has been evaluated at cosmic time t_{max}. Thus, by the time the perturbed sphere had stopped expanding, its density was already 5.55 times greater than that of the background density.

Interpreted literally, the spherical perturbed region collapsed to a black hole. It is much more likely to form some sort of bound object. The temperature of the gaseous baryonic matter increased until internal pressure gradients became sufficient to balance the attractive force of gravitation. For the cold dark matter, during collapse, the cloud fragmented into sub-units and then, through the process of *violent relaxation*, these regions came to a dynamical equilibrium under the influence of large scale gravitational potential gradients.

The Non-linear Collapse of Density Perturbations

The end result is a system which satisfies the Virial Theorem. At a_{max}, the sphere is stationary and all the energy of the system is in the form of gravitational potential energy. For a uniform sphere of radius r_{max}, the gravitational potential energy is $-3GM^2/5r_{\mathrm{max}}$. If the system does not lose mass and collapses to half this radius, its gravitational potential energy becomes $-3GM^2/(5r_{\mathrm{max}}/2)$ and, by conservation of energy, the kinetic energy, or internal thermal energy, acquired is

$$\text{Kinetic Energy} = \frac{3GM^2}{5(r_{\mathrm{max}}/2)} - \frac{3GM^2}{5r_{\mathrm{max}}} = \frac{3GM^2}{5r_{\mathrm{max}}} \,. \tag{81}$$

By collapsing by a factor of two in radius from its maximum radius of expansion, the kinetic energy, or internal thermal energy, becomes half the negative gravitational potential energy, the condition for dynamical equilibrium according to the Virial Theorem. Therefore, the density of the perturbation increased by a further factor of 8, while the background density continues to decrease.

The Non-linear Collapse of Density Perturbations

The scale factor of the perturbation reached the value $a_{\mathrm{max}}/2$ at time $t = (1.5 + \pi^{-1})t_{\mathrm{max}} = 1.81t_{\mathrm{max}}$, when the background density was a further factor of $(t/t_{\mathrm{max}})^2 = 3.3$ less than at maximum. The net result of these simple calculations is that, when the collapsing cloud became a bound virialised object, its density was $5.55 \times 8 \times 3.3 \approx 150$ times the background density at that time.

These simple calculations illustrate how structure forms according to the large scale simulations. They show that galaxies and clusters must have formed rather later than the simple estimates we gave in the second lecture. According to Coles and Lucchin, the systems become virialised at a time $t \approx 3t_{\mathrm{max}}$ when the density contrast was about 400.

Using these arguments, galaxies of mass $M \approx 10^{12}\ M_{\odot}$ could not have been virialised at redshifts greater than 10 and clusters of galaxies cannot have formed at redshifts much greater than one.

The Zeldovich Approximation

The next approximation is to assume that the perturbations were ellipsoidal with three unequal principal axes. In the *Zeldovich approximation*, the development of perturbations into the non-linear regime is followed in Lagrangian coordinates. If $\boldsymbol{x}$ and $\boldsymbol{r}$ are the proper and comoving position vectors of the particles of the fluid, the Zeldovich approximation can be written

$$\boldsymbol{x} = a(t)\boldsymbol{r} + b(t)\boldsymbol{p}(\boldsymbol{r}) \; . \tag{82}$$

The first term on the right-hand side describes the uniform expansion of the background model and the second term the perturbations of the particles' positions about the Lagrangian (or comoving) coordinate $\boldsymbol{r}$. Zeldovich showed that, in the coordinate system of the principal axes of the ellipsoid, the motion of the particles in comoving coordinates is described by a 'deformation tensor' D

$$D = \begin{bmatrix} a(t) - \alpha b(t) & 0 & 0 \\ 0 & a(t) - \beta b(t) & 0 \\ 0 & 0 & a(t) - \gamma b(t) \end{bmatrix} . \tag{83}$$

Because of conservation of mass, the density ϱ in the vicinity of any particle is

$$\varrho[a(t) - \alpha b(t)][a(t) - \beta b(t)][a(t) - \gamma b(t)] = \bar{\varrho} a^3(t) \; , \tag{84}$$

where $\bar{\varrho}$ is the mean density of matter in the Universe.

The Zeldovich Approximation

The clever aspect of the Zeldovich solution is that, although the constants α, β and γ vary from point to point in space depending upon the spectrum of the perturbations, the functions $a(t)$ and $b(t)$ are the same for all particles. In the case of the critical model, $\Omega_0 = 1$,

$$a(t) = \frac{1}{1+z} = \left(\frac{t}{t_0}\right)^{2/3} \quad \text{and} \quad b(t) = \frac{2}{5}\frac{1}{(1+z)^2} = \frac{2}{5}\left(\frac{t}{t_0}\right)^{4/3}, \tag{85}$$

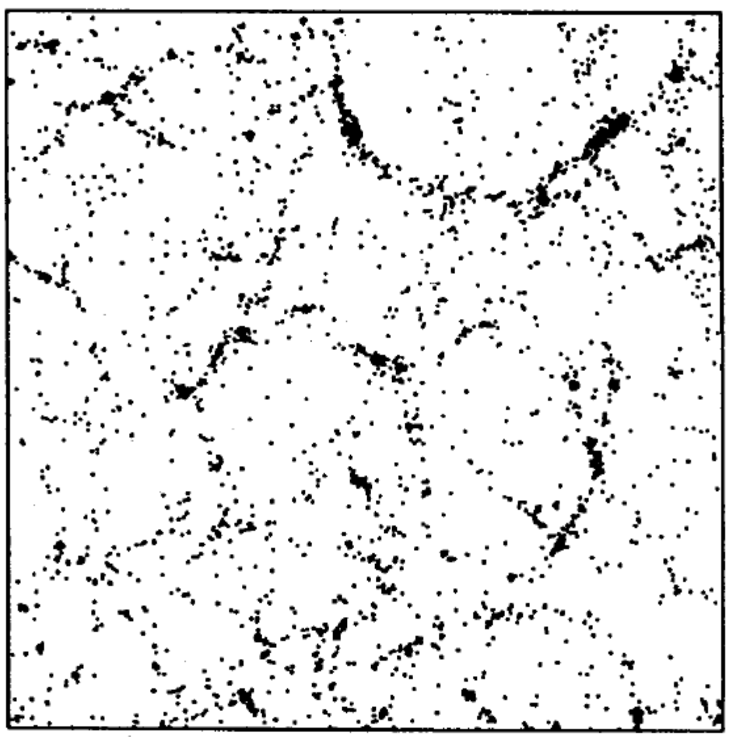
where $t_0 = 2/3H_0$. The function $b(t)$ has exactly the same dependence upon scale factor (or cosmic time) as was derived from perturbing the Friedman solutions.

If we consider the case in which $\alpha > \beta > \gamma$, collapse occurs most rapidly along the x-axis and the density becomes infinite when $a(t) - \alpha b(t) = 0$. At this point, the ellipsoid has collapsed to a 'pancake' and the solution breaks down for later times. Although the density becomes formally infinite in the pancake, the surface density remains finite, and so the solution still gives the correct result for the gravitational potential at points away from the caustic surface. The Zeldovich approximation cannot deal with the more realisitic situation in which collapse of the gas cloud into the pancake gives rise to strong shock waves, which heat the matter falling into the pancake.

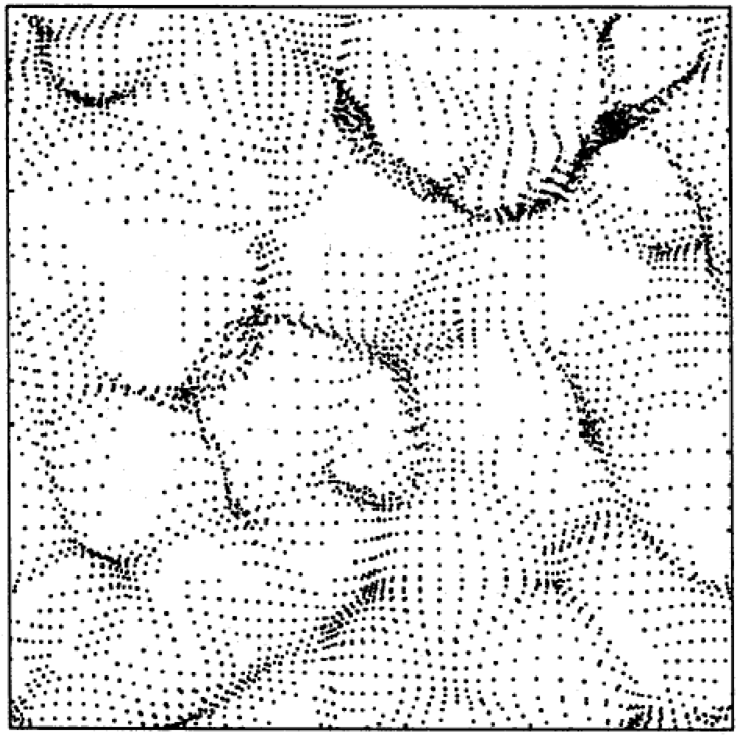
The Zeldovich Approximation

The results of numerical N-body simulations have shown that the Zeldovich approximation is quite remarkably effective in describing the evolution of the non-linear stages of the collapse of large scale structures up to the point at which caustics are formed.

N-body Simulation



Zeldovich Approximation

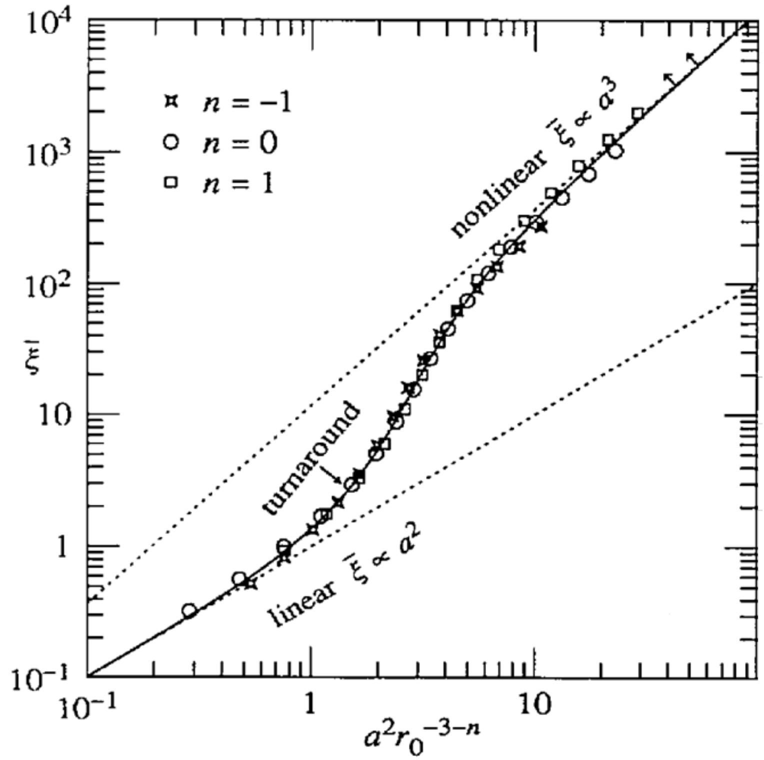


Non-Linear Development of the Density Perturbations

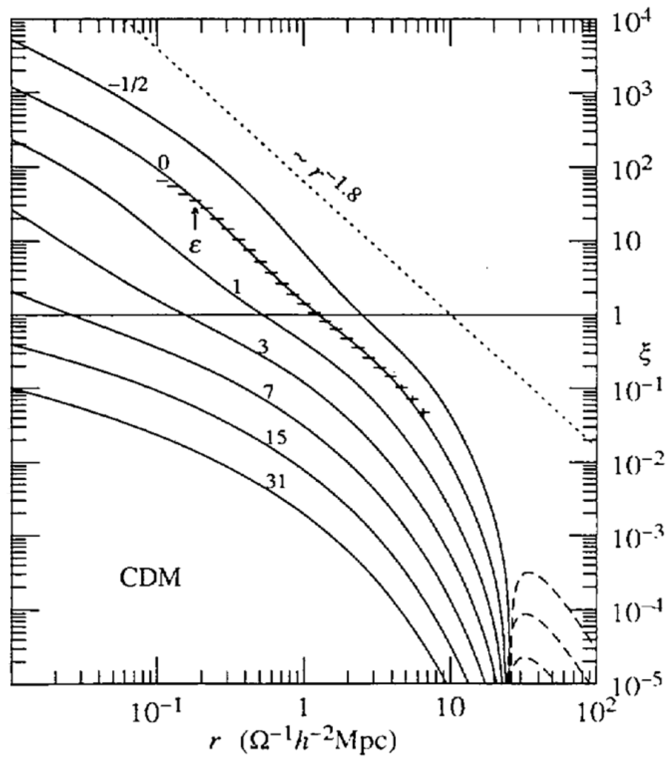
It is evident from the power-law form of the two-point correlation function for galaxies $\xi(r) = (r/r_0)^{-1.8}$ that on scales much larger than the characteristic length scale $r_0 \approx 7$ Mpc, the perturbations are still in the linear stage of development and so provide directly information about the form of the processed initial power spectrum.

On scales $r \leq r_0$, the perturbations become non-linear and it might seem more difficult to recover information about the processed power-spectrum on these scales. An important insight was provided by Hamilton and his colleagues who showed how it is possible to relate the observed spectrum of perturbations in the non-linear regime, $\xi(r) \geq 1$, to the processed initial spectrum in the linear regime.

Non-Linear Development of the Density Perturbations

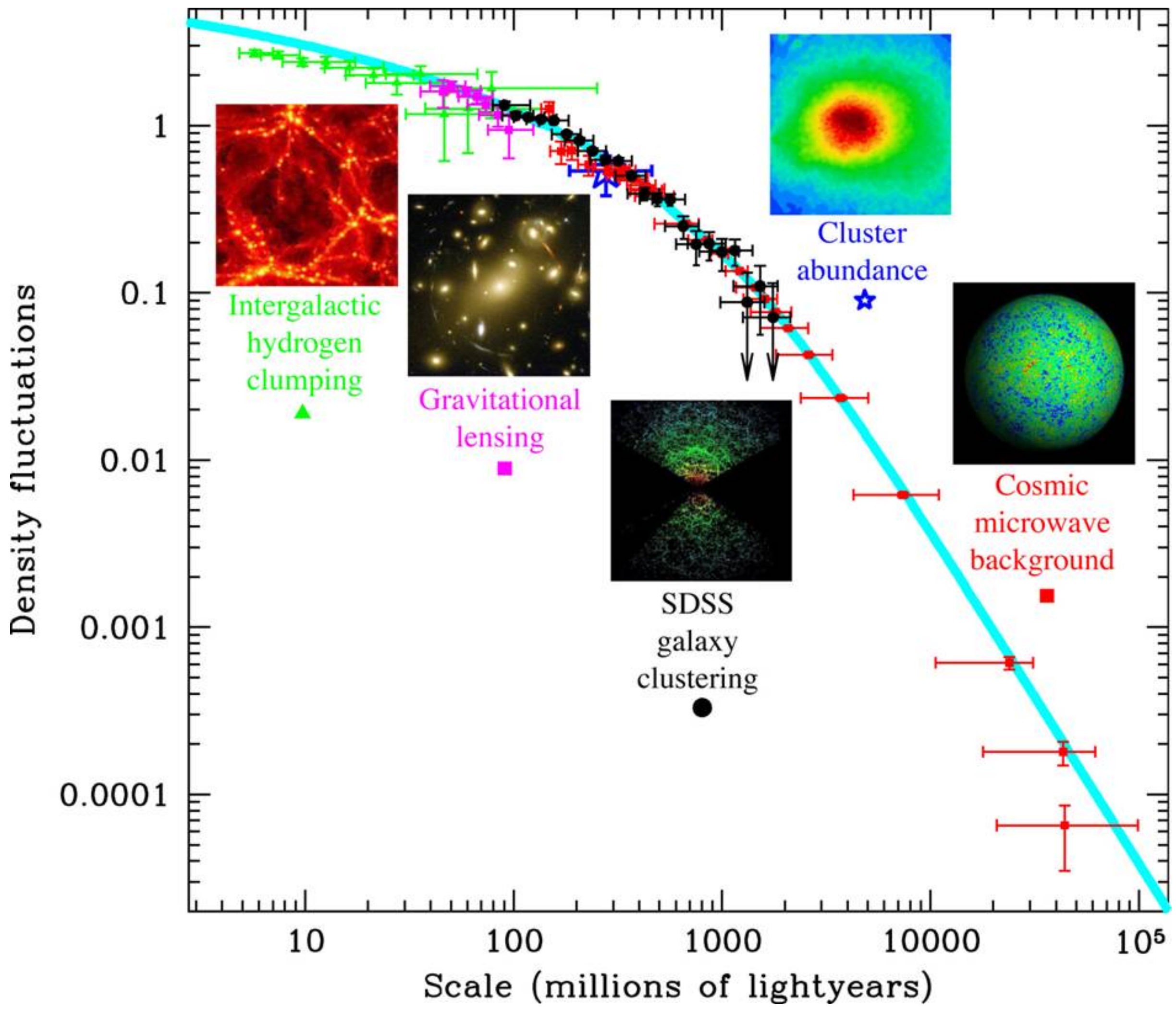


The variation of the spatial two-point correlation function with the square of the scale factor as the perturbations evolve from linear to non-linear amplitudes to bound systems.



The corresponding evolution of the spatial two-point correlation function as a function of redshift, normalised to result in a two-point correlation function for galaxies which has slope -1.8.

The Modified Initial Power Spectrum



Max Tegmark and his colleagues have shown how many other pieces of data are consistent with this picture. Note:

- Overlap of WMAP and SDSS power spectra.
- Statistics of gravitational lensing.
- Power spectrum of neutral hydrogen clouds.