



Galaxy formation and evolution

PAP 318, 5 op, autumn 2022

BK106, Exactum

**Lecture 5: Baryonic and dark matter models
of galaxy formation, 05/10/2022**



On this lecture we will discuss

1. The radiation-matter equality and cosmic sound speed in the early Universe.
2. Baryonic models of galaxy formation. Adiabatic and isothermal fluctuations in the radiation-dominated and matter-dominated eras.
3. Silk dampening of baryonic fluctuations.
4. Dark matter and galaxy formation. WIMPs as dark matter. Free-streaming of dark matter perturbations.
5. Instabilities in the presence of dark matter. The cold and hot dark matter scenarios.
6. The lecture notes correspond to: **L:** pages 350-384 (§12.4-13.8)
MBW: pages 162-176 (§4.1.2-4.1.6)



5.1 The early Universe: Radiation-matter equality

- At very high redshifts the energy density in the Universe was dominated by radiation. The redshift of the radiation-matter equality can be calculated as:

$$\frac{\rho_r}{\rho_m} = \frac{\frac{4\sigma}{c} T_0 (1+z)^4}{\Omega_0 \rho_c c^2 (1+z)^3} = \frac{2.48 \times 10^{-5} (1+z)}{\Omega_0 h^2}$$

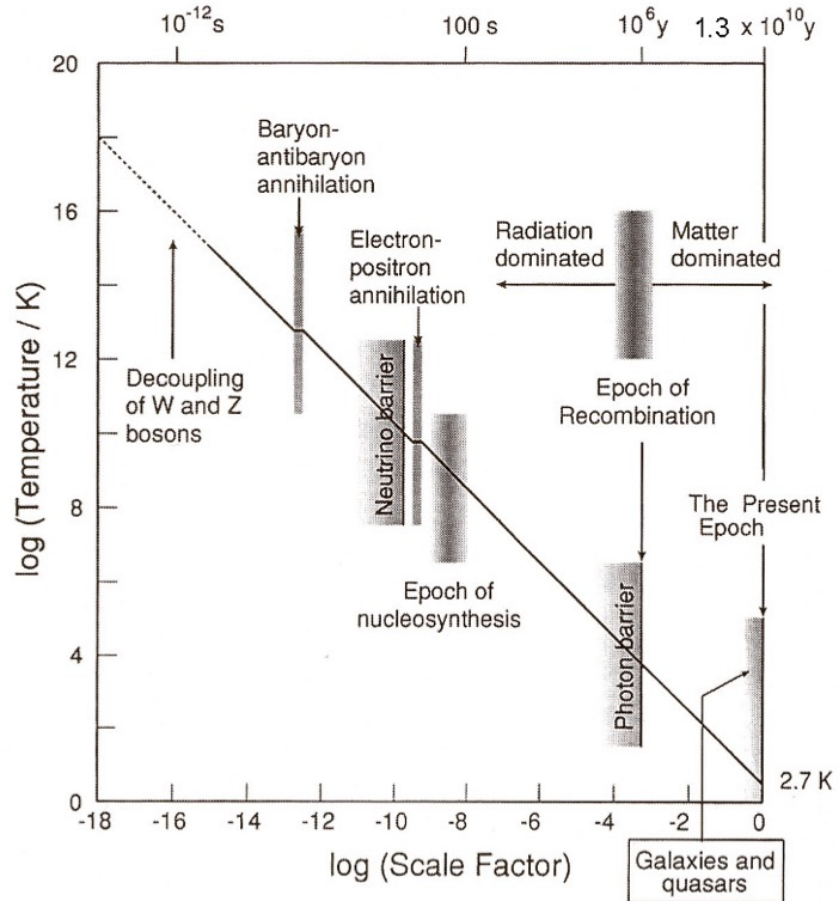
- Thus at redshifts $z \geq 4 \times 10^4 \Omega_0 h^2$, the Universe was radiation-dominated and the dynamics was described by the relation $a \propto t^{1/2}$. At lower redshifts the Universe is matter-dominated and described by the standard Friedmann models, $a \propto t^{2/3}$ provided $\Omega_0 z \gg 1$.
- The present photon-to-baryon ratio is also an important cosmological parameter, assuming $T_0 = 2.728$ K for the present-day CMB temperature:

$$\frac{N_\gamma}{N_B} = \frac{3.7 \times 10^7}{\Omega_b h^2}$$



Thermal history of the Universe

- The key epochs in the thermal history of the Universe are summarized in the figure.
- The radiation temperature decreases as $T_r \propto a^{-1}$ except for small discontinuities as different particle-antiparticle pairs annihilate at $kT \approx mc^2$.
- Key epochs: Epoch of equality of matter and radiation and epoch of recombination (formation of the CMB).
- Neutrino and photon barriers.





The sound speed as a function of cosmic epoch I

- The sound speed is proportional to the square root of the ratio of the pressure which provides the restoring force and the inertial mass density of the medium (subscript S, constant entropy, adiabatic sound waves):

$$c_S^2 = \left(\frac{\partial p}{\partial \rho} \right)_S$$

- The sound speed in the early Universe changes as the dominant contributors to the pressure, p and density ρ change. The sound speed can then be written as:

$$c_S^2 = \frac{(\partial p / \partial T)_r}{(\partial \rho / \partial T)_r + (\partial \rho / \partial T)_m} = \frac{c^2}{3} \frac{4\rho_r}{4\rho_r + 3\rho_m}$$



The sound speed as a function of cosmic epoch II

- Thus, at high redshifts, when $\rho_r \gg \rho_m$ the sound speed tends to the relativistic sound speed $c_s = c/\sqrt{3}$.
- 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 matter and radiation energy densities and the epoch of recombination, the pressure of the sound waves is provided by the radiation, but the inertia is provided by the matter

$$c_s = \left(\frac{4c^2}{9} \frac{\rho_r}{\rho_m} \right)^{1/2} = \left[\frac{16\sigma T_0^4 (1+z)}{9\Omega_m \rho_c c} \right]^{1/2} = \frac{10^6 z^{1/2}}{(\Omega_m h^2)^{1/2}}$$

- After recombination (i.e. formation of the CMB), the sound speed is the thermal sound speed of the matter, which because of the close coupling between the matter and the radiation has $T_r = T_m$.



5.2 Baryonic models of galaxy formation:

- The baryonic models provided the most natural starting point of the pioneering galaxy formation studies in the 1960s and 1970s:
 1. The Jeans' length is the maximum scale on which density perturbations can be stabilised by their internal pressure gradients.

$$\text{Matter : } \lambda_J = c_s \left(\frac{\pi}{G\rho} \right)^{1/2} \quad \text{Radiation : } \lambda_J = c_s \left(\frac{3\pi}{8G\rho} \right)^{1/2}$$

2. For wavelengths smaller than the Jeans' length, the perturbations are sound waves.
3. For wavelengths greater than the Jeans' length and for wavelengths $\lambda \gg \lambda_J$, the perturbations grow as $\Delta \propto (1+z)^{-1}$ in the matter-dominated era and as $\Delta \propto (1+z)^{-2}$ in the radiation-dominated era.
4. The horizon scale is the maximum distance over which information can be communicated at cosmic epoch t and is $r_H(t) = 3ct$ in the matter-dominated and $r_H(t) = 2ct$ in the radiation-dominated era.



Adiabatic fluctuations in the radiation-dominated era

- In the radiation-dominated era the Jeans' length was:

$$\lambda_J = \frac{c}{\sqrt{3}} \left(\frac{3\pi}{8G\rho} \right)^{1/2}$$

- The corresponding Jeans' mass was:

$$M_J = (\pi\lambda_J^3/6)\rho_B = 8.5 \times 10^{28} a^3 \Omega_B h^2 M_\odot$$

- Finally we have to consider the size of the horizon:

$$r_H = 2ct = c \left(\frac{3}{8\pi G\rho} \right)^{1/2}$$

- Thus, the Jeans length was of the same order as the horizon size. In the very early stages the perturbation is larger than the horizon scale and grows as $\Delta \propto (1+z)^{-2}$. At some stage the perturbation enters the horizon and, at more or less the same time, the Jeans length became greater than the scale of the perturbation. The perturbations were therefore stabilized against gravitational collapse at later times and became sound waves.



Adiabatic fluctuations in the matter-dominated era

- The variation of the baryonic mass within the particle horizon can be worked out as using $r_H \propto a^{3/2}$.
$$M_{b,\text{hor}} = \left(\frac{\pi r_H^3}{6} \right) \rho_b = \frac{3.0 \times 10^{22}}{(\Omega_0 h^2)^{1/2}} a^{3/2} M_\odot$$
- The Jeans' mass is now with the appropriate sound speed of $c_s = c(4\rho_{\text{rad}}/9\rho_b)^{1/2} \propto a^{-1/2}$, $\lambda_J \propto c_s \rho_b^{-1/2} \propto a$. Thus the Jeans mass will be independent of the scale factor a :
$$M_J = \frac{3.75 \times 10^{15}}{(\Omega_b h^2)^2} M_\odot$$
- Note that perturbations that are larger than this mass grew according to the standard result $\Delta \propto a$, when they came through the horizon.
- At recombination the pressure from the radiation disappears and the appropriate sound speed is now the sound speed of gas with $T=3000$ K.

$$M_J = (\pi \lambda_J^3 / 6) \rho_B = 1.6 \times 10^5 (\Omega_0 h^2)^{-1/2} M_\odot$$



The damping of sound waves I

- Although the matter and radiation are closely coupled through 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 which maintains the oscillations, the perturbation is damped out if the radiation has time to diffuse out of it. This process is referred to as Silk damping.
- At any epoch the free path of scattering of photons by electrons is $\lambda = (N_e \sigma_T)^{-1}$, where $\sigma_T = 6.665 \times 10^{-29} \text{ m}^2$ is the Thomson cross-section. According to kinetic theory the photons can diffuse over the distance:

$$r_D \approx (Dt)^{1/2} = \left(\frac{1}{3} \lambda ct\right)^{1/2}$$



The damping of sound waves II

- The Silk damping mass scale can be worked out in both the radiation-dominated epoch where $[t \propto (1+z)^{-2}]$ and the matter-dominated epoch where $[t \propto (1+z)^{-3/2}]$. Finally we use the following relation for the number density of electrons:

$$N_e = \frac{\Omega_b \rho_c (1+z)^3}{m_p} = 11 \Omega_b h^2 (1+z)^3 \text{m}^{-3}$$

- Thus we get the following expressions for the Silk damping mass in the radiation- and matter-dominated eras (see L p. 356 and the additional notes for details):

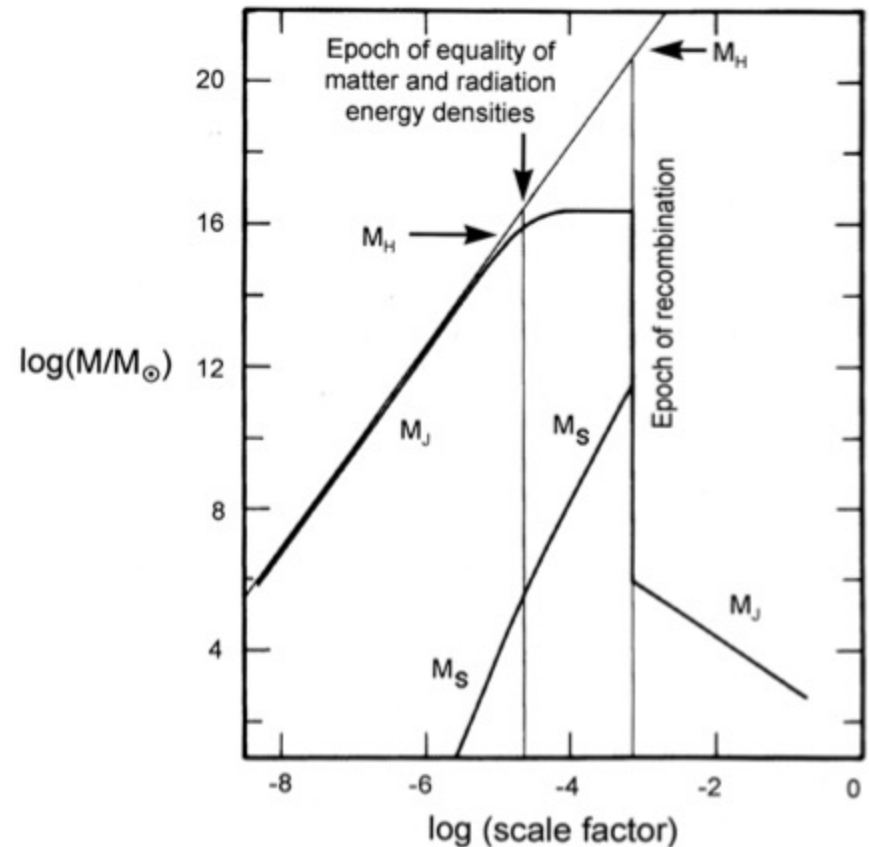
$$M_S = \frac{4\pi}{3} r_S^3 \rho_b = 2.4 \times 10^{26} (\Omega_b h^2)^{-1/2} (1+z)^{-9/2} M_\odot \quad : \text{Radiation}$$

$$M_S = \frac{4\pi}{3} r_S^3 \rho_b = 2.0 \times 10^{23} (\Omega_b h^2)^{-5/4} (1+z)^{-15/4} M_\odot \quad : \text{Matter}$$



The simple baryonic picture

- We can now put together all these ideas to develop a first picture of galaxy formation.
- This is the simplest baryonic picture and includes many features, which will reappear in the Λ CDM picture.
- The diagram shows how the horizons mass M_H , the Jeans mass M_J and the Silk dampening mass M_S evolve as a function of the scale factor a .





Isothermal perturbations

- So far we have discussed adiabatic perturbations, for which it was assumed that the density perturbations were associated with pressure perturbations according to the standard adiabatic relation:
$$\frac{\delta p}{p} = \gamma \frac{\delta \rho}{\rho}$$
- Isothermal perturbations are isothermal in the sense that, since the matter is maintained at the same temperature as the radiation by Compton scattering and the radiation is assumed to be uniform, they cause no fluctuations in the background radiation temperature during the radiation-dominated phase.
- Large-scale perturbations are similar in the adiabatic and isothermal cases, but there are major differences for small masses, since the background radiation is uniform, no Silk damping of the isothermal perturbations took place and all mass scales survived to the epoch of recombination.



The adiabatic scenario for galaxy formation

- In the adiabatic scenario developed by Yakov Zeldovich and championed by Soviet cosmologists, it was assumed that a spectrum of small adiabatic perturbations was set up in the early Universe.
- In this adiabatic picture only large-scale perturbations with masses $M \geq M_S = 10^{12} (\Omega_b h^2)^{-5/4} M_\odot$ survived, with smaller perturbations being damped out by photon diffusion (Silk damping).
- Perturbations with masses greater than the Jeans' mass $M_J = 3.75 \times 10^{15} (\Omega_b h^2)^{-2} M_\odot$ continued to grow from the time they came through the horizon.
- In this model structure formed top-down, with the largest structures with the masses of superclusters and galaxy clusters forming first. Galaxies were assumed to form through the fragmentation or thermal instabilities in the large-scale string-like structure.



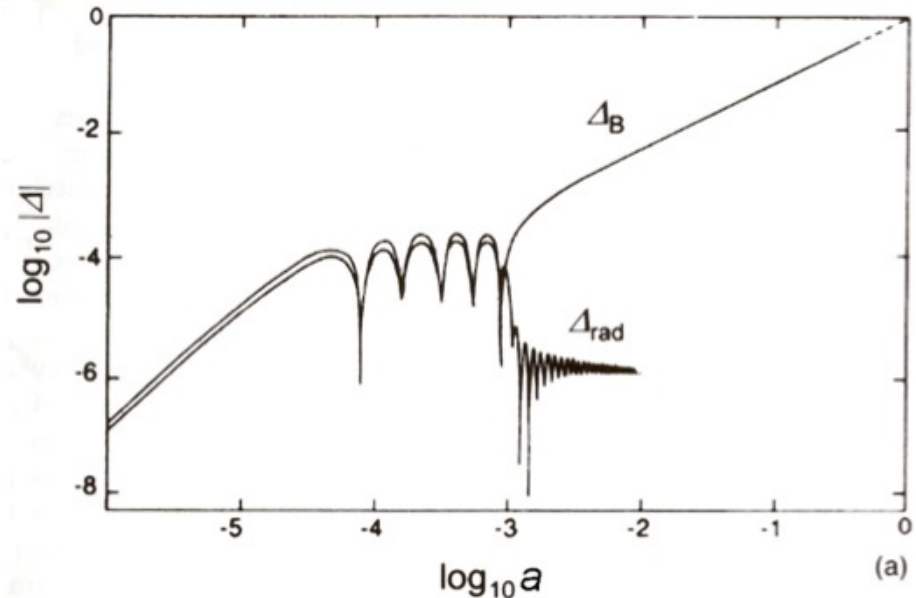
The isothermal scenario for galaxy formation

- The competing model of structure formation, the isothermal scenario developed mainly by Princeton astronomers led by Jim Peebles had the attractive feature that all perturbations larger than $M_J = 1.6 \times 10^5 (\Omega_b h^2)^{-1/2} M_\odot$ survived to the epoch of recombination and began immediately to collapse. The Jeans' mass corresponded to the mass of globular clusters, which are among the oldest objects in the Universe.
- In this model structure formation then proceeds bottom-up by hierarchical clustering, with larger objects forming through the merging of smaller-mass objects.
- When these theories were developed dark matter was not yet known and observations could not discriminate between the two models.



Problems with the baryonic models of galaxy formation

- A major problem with baryonic models of galaxy formation is that constraints from primordial nucleosynthesis show that the majority of the matter density must be in non-baryonic form.
- In purely baryonic theories, the fluctuations in the CMB had to have very large amplitudes of $\Delta \geq 3 \times 10^{-3}$, which is well above the observational constraints.



The diagram shows schematically how perturbations develop in purely baryonic structure formation models.



5.3 Dark matter and galaxy formation: Forms of dark matter

1. **Axions:** A hypothetical elementary particle that was postulated to resolve the strong CP-(charge-parity)-problem in quantum chromodynamics (QCD). If they exist, axions would be light particles with masses of 10^{-2} - 10^{-5} eV. They never acquired thermal velocities and would function as ‘cold’ dark matter.
2. **Neutrinos:** They have a small, but as of yet unknown rest mass probably in the range of ~ 0.1 eV. With this mass they contribute to the mass density of the Universe, but cannot be the dominant form of dark matter. Relativistic neutrinos function as ‘hot’ dark matter.
3. **WIMPs:** Weakly interactive massive particles have rest masses in the 1-10 GeV range. Prime candidates for WIMPs are supersymmetric particles, in particular the neutralino, which would be the lightest stable supersymmetric particle. Used to be the favoured model, but as of yet no evidence for supersymmetry has been seen at CERN.



Freeze-out of DM particles

- Initially in the early Universe particles were maintained in thermal equilibrium and the particle-antiparticle abundances were given by:

$$N_X = N_{\bar{X}} = \frac{4\pi g_X}{h^3} \int_0^\infty \frac{p^2 dp}{e^{E/kT} \pm 1}$$

- This is valid in the relativistic limit, where the “+”-sign is for fermions and “-”-sign for bosons, g_X is the statistical weight. The corresponding non-relativistic limit is:

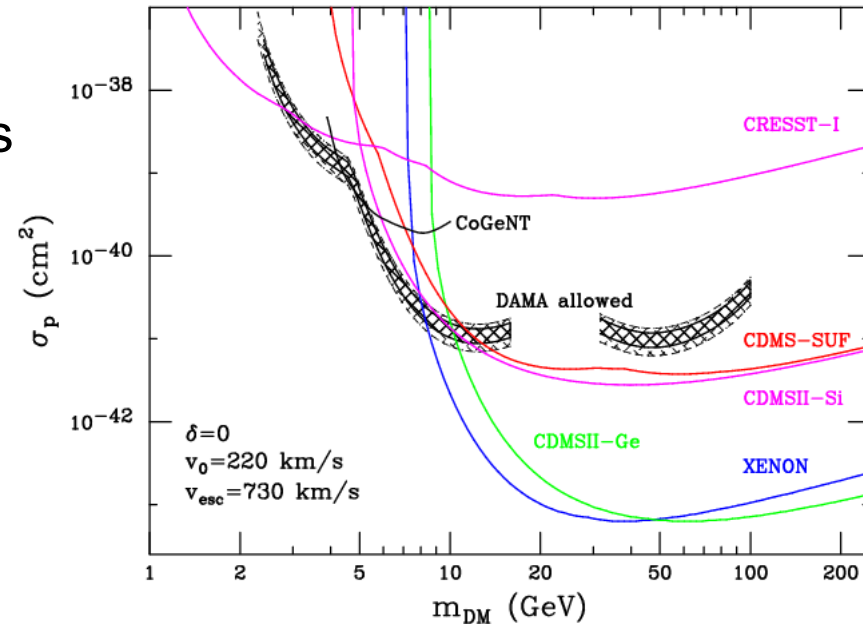
$$N_X = g_X \left(\frac{m_X kT}{h^2} \right)^{3/2} e^{-\frac{m_X c^2}{kT}}$$

- Thus, once particles became non-relativistic ($kT \ll m_X c^2$), the number densities decreased exponentially until the timescale of the interactions, which maintain species in equilibrium exceeded the expansion age of the Universe. At this point, the abundance of the massive particle froze out and remained the same if the particle was stable.



WIMPs as dark matter

- If WIMPs are the dark matter in the Universe they cannot be as common as neutrinos and photons, as their rest mass energy would exceed $\Omega_0=1$ by several orders of magnitudes.
- Thus, the decoupling of WIMPs from thermal equilibrium had to take place after they had become non-relativistic, that is after the epoch at which $kT \sim m_\chi c^2$.
- The $\Omega_0=0.3$ constraint together with the requirement of non-relativistic particles help constrain the masses and cross-sections of the dark matter particles.



Plot shows the constraints by various direct detection experiments on the dark matter particle in the mass-cross-section plane.



Metric perturbations in hot and cold dark matter

- In purely baryonic models we noted that all perturbations could be decomposed into adiabatic and isothermal modes. In dark matter models a similar analogy can be used by dividing all perturbations into curvature modes and isocurvature modes.

- The curvature modes are similar to adiabatic modes in the sense that perturbations in the radiation, baryonic matter and the dark matter were driven by gravitational perturbations in the metric:

$$\frac{1}{3} \frac{\delta\rho_b}{\rho_b} = \frac{1}{3} \frac{\delta\rho_{DM}}{\rho_{DM}} = \frac{1}{4} \frac{\delta\rho_{rad}}{\rho_{rad}} = \frac{1}{4} \frac{\delta\rho_\nu}{\rho_\nu}$$

- In isocurvature modes, the total mass-energy is constant, and so there are no perturbations in the spatial curvature of the background model, but there are fluctuations in the mass-energy density of each of the four components.



Free-streaming and the damping of hot dark matter perturbations

- As long as dark matter particles were strongly coupled in the early Universe, they behaved as ordinary relativistic or non-relativistic particles.
- However, if the particles were relativistic when they froze-out, as was the case for example with the neutrinos, they would continue to travel in ‘straight lines’ at the speed of light after this time.
- The free-streaming of the dominant dark matter particles would rapidly damp out perturbations on the horizon scale. This was not important for non-relativistic cold dark matter.
- The free-streaming mass can be calculated in the neutrino case, with the result being that all, but the the most massive perturbations would be free-streamed away: $M_{\text{FS}} \approx 4 \times 10^{15} \left(\frac{m_\nu}{30\text{eV}} \right)^{-2} M_\odot$



Instabilities in the presence of dark matter I

- 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:

$$\begin{cases} \ddot{\Delta}_B + 2 \left(\frac{\dot{a}}{a} \right) \dot{\Delta}_B = 4\pi G \rho_B \Delta_B + 4\pi G \rho_D \Delta_D \\ \ddot{\Delta}_D + 2 \left(\frac{\dot{a}}{a} \right) \dot{\Delta}_D = 4\pi G \rho_B \Delta_B + 4\pi G \rho_D \Delta_D \end{cases}$$

- Let us find the solution for the case in which the dark matter $\Omega_0=1$ and the baryon density is negligible compared with that of the dark matter. Then the second equation above 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 perturbation becomes:

$$\ddot{\Delta}_B + 2 \left(\frac{\dot{a}}{a} \right) \dot{\Delta}_B = 4\pi G \rho_D B a$$



Instabilities in the presence of dark matter II

- Since the background model is the critical model, for which $a=(3H_0t/2)^{2/3}$ and $3H_0^2=8\pi G\rho_D=8\pi G\rho_D(0)a^{-3}$, the equation simplifies to:

$$a^{3/2} \frac{d}{da} \left(a^{-1/2} \frac{d\Delta}{da} \right) + 2 \frac{d\Delta}{da} = \frac{3}{2} B$$

- The solution $\Delta=B(a-a_0)$, satisfies this equation. This means that the amplitude of the baryon fluctuations can be very small at some redshift z_0 , 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:

$$\Delta_B = \Delta_D \left(1 - \frac{z}{z_0} \right)$$

- Thus, the amplitude of the perturbations in the baryons grows rapidly to the same amplitude as the dark matter perturbations. The baryons ‘fall into’ the dark matter potential wells and rapidly attain the same amplitude.



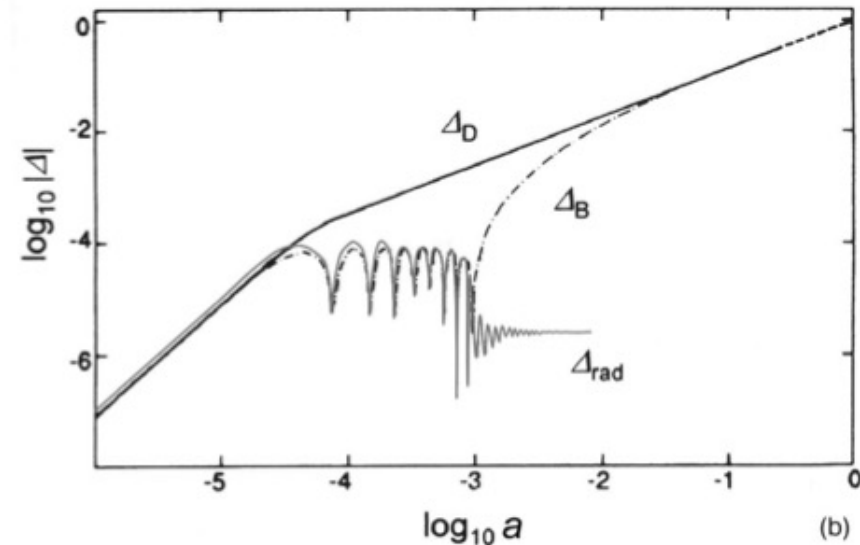
The hot dark matter (HDM) scenario

- In the hot dark matter scenario, which is now largely of historical interest the dark matter is the form of relativistic neutrinos.
- Because of free-streaming of the relativistic dark matter all perturbations that are smaller than $M_{FS} \sim 10^{15} M_{\odot}$ were wiped out, also baryonic perturbations with masses up to $M_S \sim 10^{14} M_{\odot}$ were damped out because of Silk dampening. However this is not such an important feature since the dark matter is the dominant mass component.
- After recombination the baryons fell into the dark matter pockets and structure formation could start.
- A key prediction of the HDM picture is that the first structures to form are those of the large-scale structures. Galaxies then form later through fragmentation (top-down). This poses problems for the early heating and ionisation of the intergalactic gas and the early chemical enrichment of that gas.



The standard cold dark matter (CDM) scenario I

- Separate evolution of the CDM perturbation and the coupled photon-baryon plasma from when the perturbation entered the horizon until recombination.
- Regeneration of large-amplitude perturbations in baryons after the epoch of recombination.
- The decay of perturbations in the photon gas after the epoch of recombination.
- The growth of perturbations on superhorizon scales at $a < 3 \times 10^{-5}$ as the potential fluctuations were frozen-in on super-horizon scales.



Structure development in the standard CDM picture. The observed $\Delta \sim 10^{-5}$ fluctuations in the CMB can only be reconciled with the existence of galaxies in models with cold dark matter.



The standard cold dark matter (CDM) scenario II

- The cold dark matter particles decouple when they are already non-relativistic at $t \sim 10^{-9}$ s, when the mass within the horizon is very small $M \ll M_{\odot}$.
- Free-streaming was unimportant and dark matter perturbations on all scales of astrophysical interest survived from the early Universe.
- Adiabatic baryonic fluctuations came through the horizon and were quickly stabilised by the pressure of the plasma and Silk damping of masses wiped out structures up to $M_S \sim 10^{14} M_{\odot}$.
- The CDM perturbation was decoupled from the baryons and by the time of recombination it was a factor of $z_{\text{eq}}/z_{\text{rec}}$ larger.
- After recombination the Jeans' mass dropped to $M_J \approx 7 \times 10^6 M_{\odot}$ and structure formation proceeded bottom-up.



What have we learned?

1. The sound speed changes in the early Universe depends on which components provide the pressure and density in the gas. Through the Jeans' mass it determines which structures can form.
2. In baryonic galaxy formation theories the mass of structure formation is determined by the relationship between the horizon mass, Jeans' mass and the Silk damping mass. Baryonic theories predict large fluctuations in the CMB of the order of $\delta T/T \sim 10^{-3}$.
3. WIMPs are the leading candidate for cold dark matter. In these models the CDM perturbation is decoupled from the baryon-radiation mix and after recombination the baryons can fall into the potential well created by the DM, structures form bottom-up.
4. In hot dark matter (neutrino) models, the relativistic dark matter streams freely wiping out structures on scales below $M_{FS} \sim 10^{15} M_{\odot}$, with the largest structures forming first (top-down).