Part III Cosmology

James Fergusson

Contents

1 Geometry and Dynamics 7

1.1 Metric 7
1.2 Kinematics 9
1.3 Redshift 10
1.4 Dynamics 11

2 Inflation 19

2.1 Horizon Problem 19
2.2 Horizon Solution 21
2.3 Physics of inflation 23
2.4 Reheating* 26
2.5 Issues with inflation? 27

3 Cosmological Perturbation Theory 33

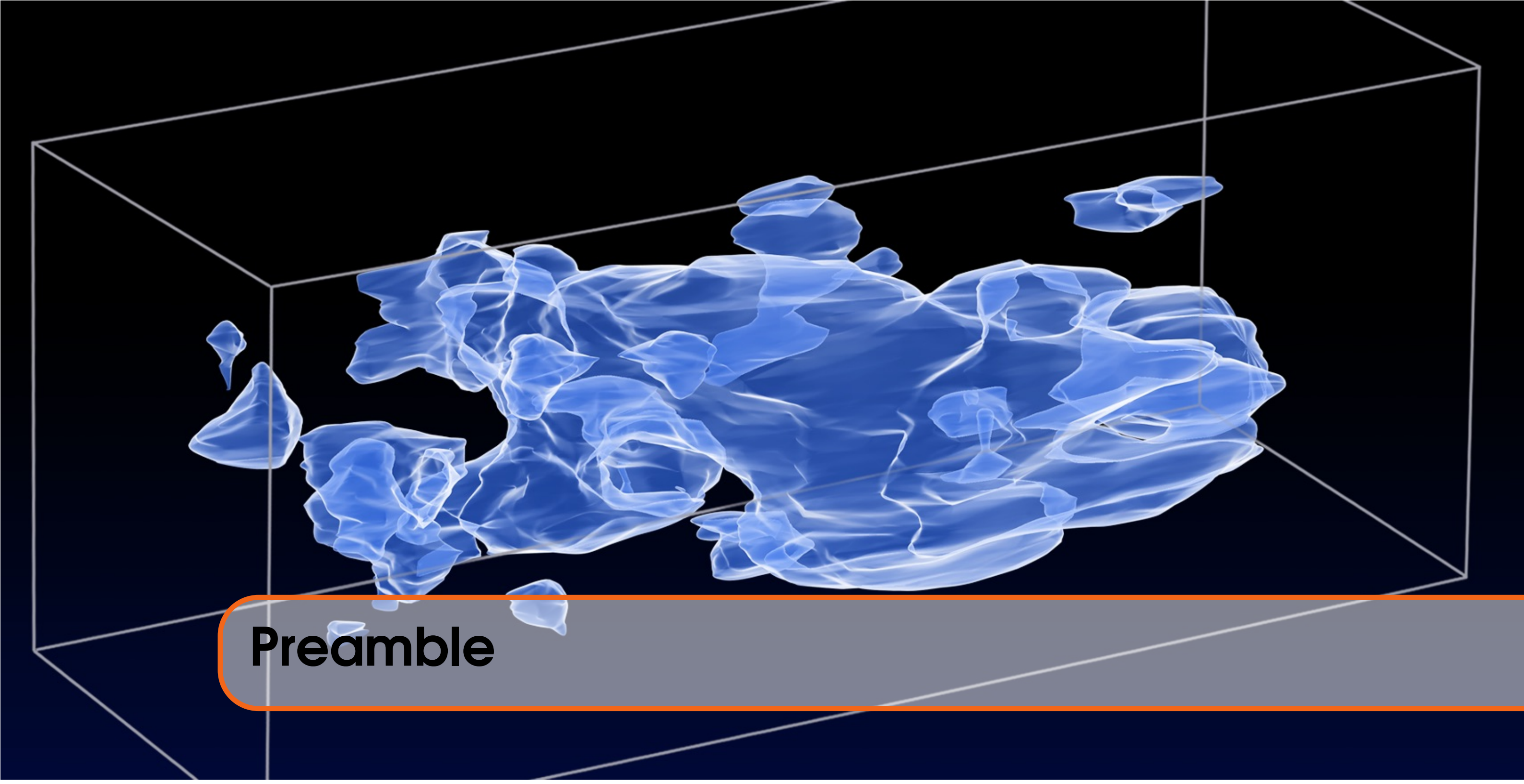
3.1 Perturbed metric 33
3.2 Gauge Problem 34
3.3 Perturbed Matter 36
3.4 Linearised Equations 37
3.5 Adiabatic Perturbations 40
3.6 Curvature Perturbation 40

4 Structure Formation - Linear 47

4.1 Initial Conditions 48

4.2 Potential ϕ 49
4.2.1 Radiation Domination 49
4.2.2 Matter Domination 49
4.3 Radiation Perturbation 49
4.3.1 Radiation Domination 49
4.3.2 Matter Domination 50
4.4 Matter Perturbation 51
4.4.1 Matter Domination 51
4.4.2 Radiation + Matter Eras 51
4.5 Baryon Perturbation 53
4.5.1 Radiation Domination 53
4.5.2 Matter Domination 53
4.5.3 Baryon Acoustic Oscillations 54
4.6 Power Spectra 56
4.6.1 Initial Conditions 56
4.6.2 Matter Power Spectrum 56
4.6.3 Radiation Power Spectrum 57

5 Structure Formation - Halos 61
5.1 Spherical Collapse 61
5.2 Filtering 64
5.3 Extended Press-Schechter Mass Function 66
5.4 Linear Bias 70



Preamble

Details:

Dr James Fergusson
Room: B1:12
J.Fergusson@DAMTP.cam.ac.uk

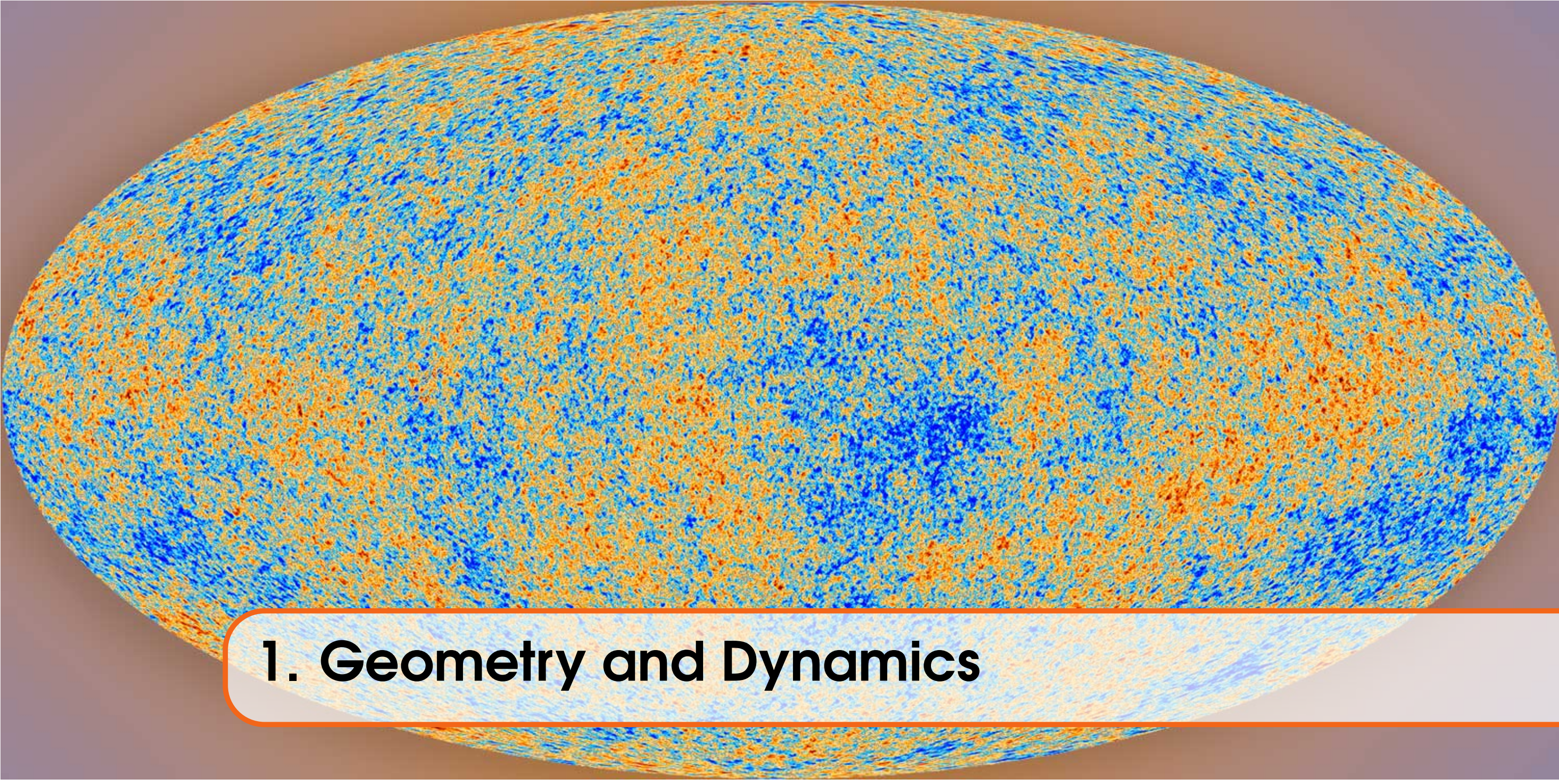
Textbooks:

Modern Cosmology - Dodelson
The Early Universe - Kolb, Turner
Cosmology - Weinberg

Online notes:

Daniel Baumann - http://www.damtp.cam.ac.uk/user/db275/Cosmology.pdf
Anthony Lewis - http://cosmologist.info/teaching/EU/
Antonio Riotto - http://www-ucjf.troja.mff.cuni.cz/malinsky/cejp2013/notes.pdf
Christian Knobel - ArXiv:12085931 [astro-ph.CO]

For those new to GR   Sean Carroll - https://preposterousuniverse.com/wp-content/uploads/2015/08/grtinypdf.pdf



1. Geometry and Dynamics

Conventions:

We will use the following conventions for GR:

1. Metric signature will be (-,+,+,+)
2. 4-Space indices will be Greek $(\mu, \nu, \dots)$ 3-space indices will be Latin $(i, j, \dots)$
3. index 0 will refer to space and (1,2,3) will be space

We will also use natural units where $c = \hbar = 1$

1.1 Metric

We will make the following assumptions regarding the universe on large scales ($\geq 100\,\text{Mpc}$):

1. The universe is Isotropic / Same in all directions / Rotationally invariant
2. The universe is Homogeneous / Same at all locations / Translationally invariant
3. General Relativity is the correct description on these scales and densities

The first two assumptions have strong support from the homogeneity and isotropy of the Cosmic Microwave Background (CMB) whose temperature is uniform to one part in 100,000 and the tiny fluctuations in temperature are statistically isotropic.

The third has strong support on solar system scales and strong support on larger scales but only if we include “Dark Matter" and “Dark Energy" which do not exist in the standard model (the following article sums up the evidence for GR on different scales ArXiv:1412.3455)

These three assumptions force the metric to take the simple form

$$ds^2 = g_{\mu\nu} dX^\mu dX^\nu = -dt^2 + a^2(t) d\ell^2 \tag{1.1}$$

where

$$d\ell^2 = \gamma_{ij} dx^i dx^j \tag{1.2}$$

is a constant curvature 3-metric which can take three forms:

1. Positive curvature / Spherical ($\mathbb{S}^3$) / $d\ell^2 = d\mathbf{x}^2 + du^2$ with $|\mathbf{x}|^2 + u^2 = R^2$

2. Zero curvature / Euclidean ($\mathbb{E}^3$) / $d\ell^2 = d\mathbf{x}^2$
3. Negative curvature / Hyperbolic ($\mathbb{H}^3$) / $d\ell^2 = d\mathbf{x}^2 - du^2$ with $|\mathbf{x}|^2 - u^2 = -R^2$

Aside: To see why this is correct we can think of embedding a 2-sphere ($\mathbb{S}^2$) into a Euclidian 3-space ($\mathbb{E}^3$) (unfortunately it's not possible to do the same with hyperbolic spaces so you'll just have to accept the analogy). If we use cylindrical coordinates then the 3-metric is

$$d\ell^2 = dr^2 + r^2 d\theta^2 + dz^2 = d\mathbf{x}^2 + dz^2 \tag{1.3}$$

with the equation of a sphere centred at the origin with radius R

$$r^2 + z^2 = |\mathbf{x}|^2 + z^2 = R^2 \tag{1.4}$$

So the equations combined describe a space with constant curvature $1/R^2$

Now we can eliminate the dummy variable u to obtain the correct form for the 3-metric. First we will rescale the coordinates so $\mathbf{x} \to R\mathbf{x}$ and $u \to Ru$ so that

$$d\ell^2 = R^2\left(d\mathbf{x}^2 \pm du^2\right), \quad \mathbf{x}^2 \pm u^2 = \pm 1 \tag{1.5}$$

Now taking the constrain we can deduce $udu = \mp \mathbf{x} d\mathbf{x}$ and substitution gives

$$\begin{aligned} d\ell^2 &= R^2\left(d\mathbf{x}^2 + K\frac{(\mathbf{x}\cdot d\mathbf{x})^2}{1-K\mathbf{x}^2}\right) \\ &= R^2\left(\frac{dr^2}{1-Kr^2} + r^2 d\Omega^2\right), \quad K = \begin{cases} 1 & \mathbb{S}^3 \\ 0 & \mathbb{E}^3 \\ -1 & \mathbb{H}^3 \end{cases} \end{aligned} \tag{1.6}$$

where we have introduced K, the comoving curvature, to take care of the three topologies. Finally we can substitute the 3-metric back into Eq.1.1 and absorb the radius of curvature of the 3-metric into the scale factor, $a(t)R \to a(t)$ to obtain the Robertson-Walker metric (or the Freidmann-RW metric or F-Lemaitre-RW metric depending on your generosity)

FRW Metric:

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-Kr^2} + r^2 d\Omega^2\right) \tag{FRW}$$

Notes:
1. The metric is invariant under rescaling $a \to \lambda a$, $r \to r/\lambda$, and $K \to K/\lambda^2$which we can use to set $a_0 = 1$ today. However this means that $k \neq \pm 1$ anymore
2. The spatial metric is in comoving coordinates. These are related to physical coordinates by $r_{ph} = ar$ and $K_{ph} = K/a^2 = 1/R_{ph}^2$ where R_{ph} is the physical radius of curvature. The physical velocity is

   $$v_{ph} = \frac{dr_{ph}}{dt} = a\frac{dr}{dt} + \dot{a}r = v_{pec} + Hr_{ph} \tag{1.7}$$

   so is a combination of the peculiar velocity and the Hubble flow due to expansion. We have also defined the Hubble parameter as

   $$\boxed{H = \frac{\dot{a}}{a}} \tag{1.8}$$

3. It is common to use conformal time rather than proper time where the two are related by $dt = ad\tau$ and so $\tau = \int dt/a$. Then the metric takes the form

$$ds^2 = a^2(\tau)\left(-d\tau^2 + \frac{dr^2}{1-Kr^2} + r^2 d\Omega^2\right) \tag{1.9}$$

and the metric in the bracket is now static.

4. Some people prefer to use $\chi = \int dr/\sqrt{1-Kr^2}$ so the metric takes the form

$$ds^2 = a^2(\tau)\left(-d\tau^2 + d\chi^2 + \begin{cases}\sin^2\chi \\ \chi^2 \\ \sinh^2\chi\end{cases} d\Omega^2\right) \tag{1.10}$$

1.2 Kinematics

If gravity is the only force to act on a particle then it will move along a geodesic. Suppose we have a particle mass m with 4 velocity

$$U^\mu \equiv \frac{dX^\mu}{ds} \tag{1.11}$$

Then the motion of the particle is defined by the geodesic equation

$$\frac{dU^\mu}{ds} + \Gamma^\mu_{\alpha\beta}U^\alpha U^\beta = 0 \tag{1.12}$$

where we have used the Christoffel symbol defined by

$$\Gamma^\mu_{\alpha\beta} \equiv \frac{1}{2}g^{\mu\lambda}\left(\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\lambda} - \partial_\lambda g_{\alpha\beta}\right) \tag{1.13}$$

Now U^μ is the tangent vector to the geodesic so we can promote it to a vector field in the neighbourhood of the particle and write

$$\frac{dU^\mu}{ds} = \frac{dX^\alpha}{ds}\frac{dU^\mu}{dX^\alpha} = U^a\partial_\alpha U^\mu \tag{1.14}$$

And the geodesic equation becomes

$$U^\alpha\left(\partial_\alpha U^\mu + \Gamma^\mu_{\alpha\beta}U^\beta\right) = U^\alpha\nabla_\alpha U^\mu = 0 \tag{1.15}$$

where ∇_α is the covariant derivative defined by the above. In terms of momentum, $P^\mu = mU^\mu$ we have

$$\boxed{P^\alpha\partial_\alpha P^\mu = -\Gamma^\mu_{\alpha\beta}P^\alpha P^\beta} \tag{1.16}$$

which works for massless particles too!

Exercise: Show that for the FRW metric where

$$g_{00} = -1\,, \quad g_{0i} = 0\,, \quad g_{ij} = a^2(t)\gamma_{ij} \tag{1.17}$$

that

$$\Gamma^{\mu}_{00} = \Gamma^{0}_{0\beta} = 0 \tag{1.18}$$

$$\Gamma^{0}_{ij} = a\dot{a}\gamma_{ij} \tag{1.19}$$

$$\Gamma^{i}_{0j} = H\delta^{i}_{j} \tag{1.20}$$

$$\Gamma^{i}_{jk} = \frac{1}{2}\gamma^{il}\left(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}\right)\left(= 0\,\text{for}\,\mathbb{E}^3\right) \tag{1.21}$$

Now homogeneity and isotropy demand that $\partial_i P^{\mu} = 0$ so we have

$$P^0\frac{P^{\mu}}{dt} = -\Gamma^{\mu}_{\alpha\beta}P^{\alpha}P^{\beta} = -\left(2\Gamma^{\mu}_{0j}P^0 + \Gamma^{\mu}_{ij}P^i\right)P^j \tag{1.22}$$

Notes:
1. If $P^i = 0$ then $\partial_t P^i = 0$ so particles at rest remain so
2. The $\mu = 0$ part gives

$$E\frac{dE}{dt} = -\frac{\dot{a}}{a}p^2, \tag{1.23}$$

where $p^2 \equiv a^2\gamma_{ij}P^iP^j$ is the physical 3-momentum. Now as $-m^2 = -E^2 + p^2$ we have $EdE = pdp$ so

$$\frac{\dot{p}}{p} = -\frac{\dot{a}}{a} \implies p(t) \propto \frac{1}{a(t)} \tag{1.24}$$

and we see that 3-momentum decays with the expansion of the universe. For massless particles $p = E$ so energy decays and for massive particles $p = mv/\sqrt{1-v^2}$ so peculiar velocities decay

1.3 Redshift

One key effect of the fact that 3-momentum decays is that light from distant Galaxies is redshifted.

For light $\lambda - h/p$ and $p \propto 1/a$ which implies that the wavelength evolves as $\lambda \propto a$. If we consider a photon emitted at time t_1 with wavelength λ_1 and observed at time t_0 with wavelength

$$\lambda_0 = \frac{a_0}{a_1}\lambda_1 > \lambda_1 \tag{1.25}$$

we define redshift as the fractional change in the wavelength

$$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1} \tag{1.26}$$

and if we rescale a such that $a_0 = 1$ then we have

$$1 + z = \frac{1}{a} \tag{1.27}$$

For nearby sources we can expand $a(t)$ around t_0

$$a = \left(1 - (t_0 - t_1)H_0 - \frac{1}{2}H_0^2 q_0 (t_0 - t_1)^2 + \dots\right) \tag{1.28}$$

Where H_0 is the Hubble parameter today and q_0 is the deceleration parameter. Now for light, distance is $d = t_0 - t_1$ (setting $c = 1$) so to first order we can write

$$\boxed{z = H_0 d + \dots} \tag{1.29}$$

which is Hubble's law (see Fig.[1.1])

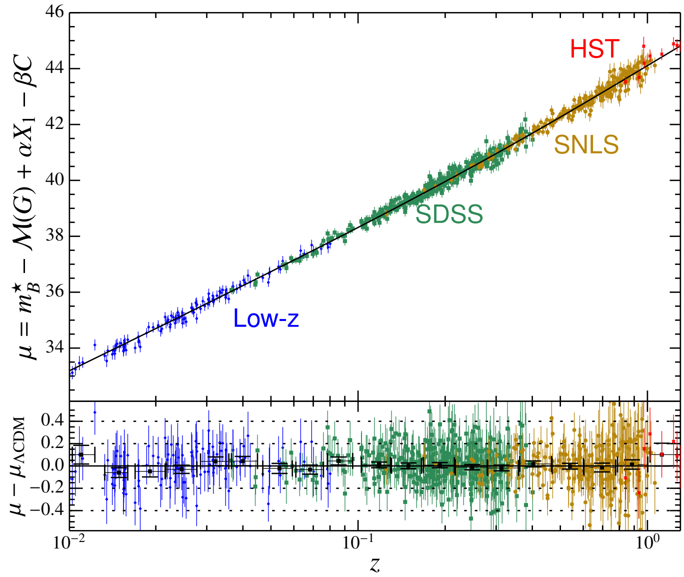


Figure 1.1: Hubble plot from type 1A supernova. μ is proportional to $\log(d_L)$

1.4 Dynamics

The only free function we have in the metric is $a(t)$ so it contains all information on the evolution of the background universe. To determine its evolution we need the Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{1.30}$$

where $T_{\mu\nu}$ is the Energy-Momentum tensor

$$T_{\mu\nu} = \begin{bmatrix} T_{00} & T_{0i} \\ T_{j0} & T_{ij} \end{bmatrix} = \begin{bmatrix} \text{Energy Density} & \text{Energy Flux} \\ \text{Momentum Density} & \text{Stress Tensor} \end{bmatrix} \tag{1.31}$$

For a comoving observer our assumptions of homogeneity and isotropy severely limit the form $T_{\mu\nu}$ can take. The most general form is

$$T_{00} = \rho(t)\,, \quad T_{0i} = 0\,, \quad T_{ij} = P(t)g_{ij} \tag{1.32}$$

which is the Energy-Momentum tensor for a perfect fluid. For a general observer the tensor takes the form

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu + P g_{\mu\nu} \tag{1.33}$$

Aside: The Einstein equations can be written with a cosmological constant Λ

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{1.34}$$

but here we will always treat is as dark energy, absorbing it into the Energy-Momentum tensor as a fluid with

$$\rho = -P \equiv \frac{\Lambda}{8\pi G} \tag{1.35}$$

This is purely semantic, and results are equivalent in either interpretation, but here we follow the idea that it is a fluid rather than an arbitrary free constant in the theory.

The Energy-Momentum tensor is conserved so we can write

$$\nabla_\mu T^\mu_\nu = \partial_\mu T^\mu_\nu + \Gamma^\mu_{\mu\lambda} T^\lambda_\nu - \Gamma^\lambda_{\mu\nu} T^\mu_\lambda = 0 \tag{1.36}$$

which implies a conservation law for the fluid

$$\boxed{\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0} \tag{1.37}$$

This can be re-written as

$$\frac{d(\rho a^3)}{dt} = -P\frac{d(a^3)}{dt} \implies dU = -PdV \tag{1.38}$$

so is simply the first law of thermodynamics (the expansion is adiabatic so $dS = 0$). If we assume a constant equation of state

$$\boxed{\omega = \frac{P}{\rho}} \tag{1.39}$$

then we can write

$$\frac{\dot{\rho}}{\rho} = -3(1+\omega)\frac{\dot{a}}{a} \implies \rho = \rho_0 \left(\frac{a}{a_0}\right)^{-3(1+\omega)} \tag{1.40}$$

All known particles have one of three equations of state

Cosmic Inventory

Matter (m)	-Cold Dark Matter (c) -Baryons (b)	$\omega = 0$	$\rho \propto a^{-3}$
Radiation (m)	-Photons(γ) -Neutrinos (ν) -Gravitons (g)	$\omega = 1/3$	$\rho \propto a^{-4}$
Dark Energy (Λ)	Vaccum Energy or Modified Gravity or ????	$\omega = -1$	$\rho \propto a^0$ = const.

Now for the curvature. We have the following definitions:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \tag{1.41}$$

$$R_{\mu\nu} \equiv \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\rho} \tag{1.42}$$

$$R \equiv g^{\mu\nu} R_{\mu\nu} \tag{1.43}$$

Exercise: Show

$$R_{00} = -3\frac{\ddot{a}}{a} \tag{1.44}$$

$$R_{0i} = 0 \tag{1.45}$$

$$R_{ij} = \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] g_{ij} \tag{1.46}$$

$$G_{00} = 3\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] \tag{1.47}$$

$$G_{0i} = 0 \tag{1.48}$$

$$G_{ij} = -\left[2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] g_{ij} \tag{1.49}$$

Note: both $R_{\mu\nu}$ and $G_{\mu\nu}$ have the same form, as required by homogeneity and isotropy, that we saw for the EM-tensor, ie: $X_{00} = A(t)\,, X_{0i} = 0\,, X_{ij} = B(t)g_{ij}$

Now we can use the Einstein equations to derive

$$G_{00} = 8\pi G T_{00} \implies 3\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right] = 8\pi G\rho \tag{1.50}$$

$$G_{ij} = 8\pi G T_{ij} \implies -\left[2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right] = 8\pi G P \tag{1.51}$$

which we can re-arrange to the Friedmann equations:

Friedmann Equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{F1}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{F2}$$

$$(\dot{F1}) + (F2) \implies \dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) \tag{F3}$$

which are the key equations governing the evolution of our universe. One subtlety is that (F3) holds for each component separately as their E-M tensors are separately conserved. (F1) and (F2) only hold for $\rho = \sum_i \rho_i$ and $P = \sum_i P_i$.

Now we need to see what solutions they allow. If we take (F1), using $\dot{a}/a = H$, we have

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{1.52}$$

and we can see that $K = 0$ defines a critical density for a flat universe

$$\boxed{\rho_{crit,0} = \frac{3H_0^2}{8\pi G}} \tag{1.53}$$

where subscript 0 denotes the quantity is evaluated today. Currently $\rho_{crit,0} \approx 8 \times 10^{-26}\mathrm{g\,cm}^{-3}$. We can use this to define the fractional density

$$\Omega_X = \frac{\rho_X}{\rho_{crit,0}} \tag{1.54}$$

so now $\sum_i \Omega_i = 1$ defines a flat universe. We can calculate the fractional density for each fluid component to deduce

$$\Omega_r = \frac{8\pi G}{3H_0^2}\rho_r = \frac{\Omega_{r,0}}{a^4} \tag{1.55}$$

$$\Omega_m = \frac{8\pi G}{3H_0^2}\rho_m = \frac{\Omega_{m,0}}{a^3} \tag{1.56}$$

$$\Omega_\Lambda = \Omega_{\Lambda,0} \tag{1.57}$$

where we have set $a_0 = 1$. Using these we can write the Friedmann equations in an convenient form

$$\dot{a}^2 = H_0^2\left(\frac{\Omega_{r,0}}{a^2} + \frac{\Omega_{m,0}}{a} + \Omega_{\Lambda,0}a^2\right) - K \tag{1.58}$$

$$\ddot{a} = -H_0^2\left(\frac{\Omega_{r,0}}{a^3} + \frac{\Omega_{m,0}}{2a^2} - \Omega_{\Lambda,0}a\right) \tag{1.59}$$

Firstly these show that for the majority of the evolution the universe the equations are dominated by a single fluid

1. Early times - **Radiation Domination (RD)** - $0 < a < \Omega_{r,0}/\Omega_{m,0}$
2. Intermediate times - **Matter Domination (MD)** - $\Omega_{r,0}/\Omega_{m,0} < a < (\Omega_{m,0}/\Omega_{\Lambda,0})^{1/3}$
3. Late times - **Dark Energy Domination (ΛD)** - $(\Omega_{m,0}/\Omega_{\Lambda,0})^{1/3} < a$

Secondly we can see from (1.58) that in order to have $\dot{a}^2 \geq 0$ for a given set of fractional densities we have a maximum allowed curvature K. We can then use this to deduce the possible solutions for the trajectories of a by considering a plot of K versus a. Let's consider a universe filled with ordinary matter first so we will set $\Omega_\Lambda = 0$. K is constant so trajectories for a correspond to horizontal lines on the plot. As $\Omega_\Lambda = 0$ we can see from (1.59) that $\ddot{a} < 0$ always and so the curve $\dot{a}^2 = 0$ represents a maximum (for a). Now from the Figure 1.2 we see that there are only 3 qualitatively different trajectories:

1. $k > 0$ so we have $\mathbb{S}^3$ topology. We expand from $a = 0$ until hitting the red no-go region which, as discussed, must be a maximum, then recollapse back to $a = 0$. This is known as a **Closed** universe.
2. $k = 0$ so we have $\mathbb{E}^3$ topology. We start at $a = 0$ and expand forever but $\dot{a} \to 0$ as $t \to \infty$. This is known as a **Flat** universe.
3. $k < 0$ so we have $\mathbb{H}^3$ topology. We start at $a = 0$ and expand forever but $\dot{a} \to \sqrt{-K}$ as $t \to \infty$. This is known as a **Open** universe.

Aside: The plots are plots of $\mathrm{sign}(K)\log|K|$ vs $\log a$ therefore we have cut the upper $\mathbb{S}^3$ region and lower $\mathbb{H}^3$ region plots at $\log|K| = 0$ and stitched them together with the $\mathbb{E}^3$ ($k = 0$) case. This is to allow the $\dot{a}^2 = 0$ curves for each fluid to be visible in the $\mathbb{S}^3$ region (and to be straight lines). I hope this does not confuse anyone as the curves appear to intersect the axis which in reality they never reach.

This clearly allows no static solutions as $\ddot{a} \neq 0$ at any location which upset Einstein as the time. His solution was to add some Dark Energy (Cosmological Constant). Setting $\Omega_\Lambda > 0$ our

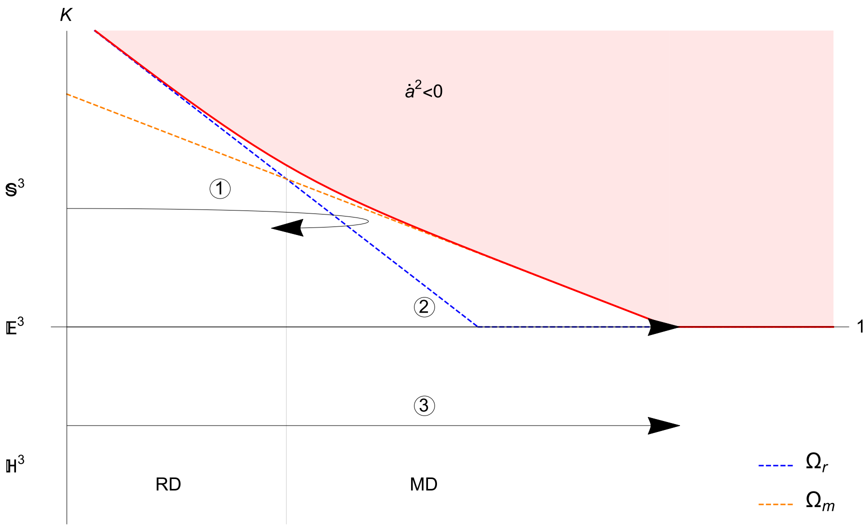


Figure 1.2: Plot of evolution of a for given curvatures without dark energy. We see there are only three possible trajectories

plot changes to Figure 1.3 and we have additional possibilities. Now during radiation or matter domination $\ddot{a} < 0$ but during Λ domination $\ddot{a} > 0$ so the universe is accelerating. We now have 5 qualitatively different trajectories and one static solution

1. $k > 0$ so we have $\mathbb{S}^3$ topology. We expand from $a = 0$ until hitting the red no-go region, which has $\ddot{a} < 0$, so must be a maximum, then recollapse back to $a = 0$. This is known as a **Closed** universe.
2. $k > 0$ so we have $\mathbb{S}^3$ topology. We collapse from $a = \infty$ until hitting the red no-go region, which has $\ddot{a} > 0$, so must be a minimum, then expand back to $a = \infty$. This is known as a **Bouncing** universe.

ESU. $k > 0$ so we have $\mathbb{S}^3$ topology. This solution has $\dot{a} = \ddot{a} = 0$ and so is static. this is the famous **Einstein Static Universe**

3. $k > 0$ so we have $\mathbb{S}^3$ topology. We expand from $a = 0$ passing arbitrarily close to the ESU static solution where expansion stalls, we eventually then move past and undergo accelerating expansion forever. This is known as a **Loitering** universe.
4. $k = 0$ so we have $\mathbb{E}^3$ topology. We start at $a = 0$ and expand forever eventually entering an accelerating phase at late times. This is known as a **Flat** universe
5. $k < 0$ so we have $\mathbb{H}^3$ topology. We start at $a = 0$ and expand forever eventually entering an accelerating phase at late times. This is known as a **Open** universe

As you will see on the example sheet the Einstein Static Universe is unstable so the main result of this analysis is that we are forced to have a dynamic universe. And to have a period of matter domination, so structure can form, it must be one that began at $a = 0$. We must live in a universe which began with a **Big Bang**.

Observations of our universe from CMB + Supernova + BAO constrain our universe to contain

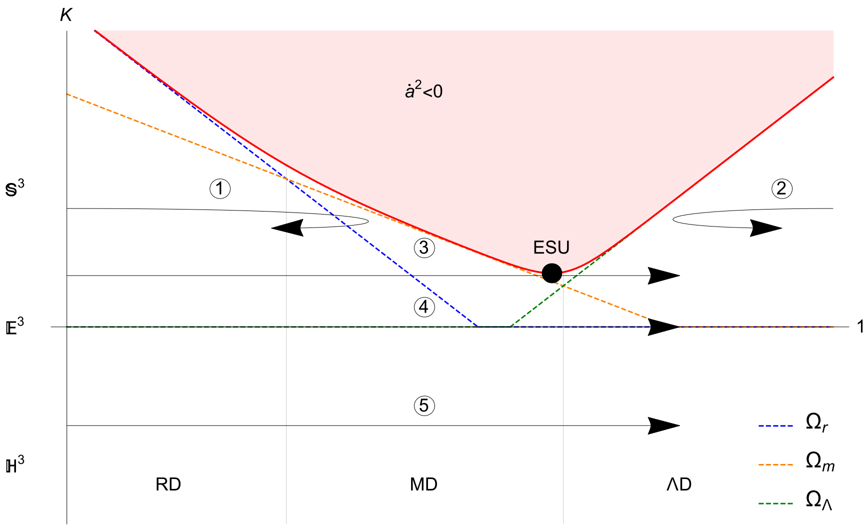


Figure 1.3: Plot of evolution of a for given curvatures with dark energy. Now we see there are five possible trajectories and one static solution

fluids in the following proportions:

$$\begin{aligned} \Omega_m = 0.32 \rightarrow &\begin{cases} \Omega_b = 0.05 \\ \Omega_c = 0.27 \end{cases} \\ \Omega_\gamma &= 1 \times 10^{-4} \\ 1 \times 10^{-3} < \Omega_\nu &< 2 \times 10^{-2} \\ \Omega_\Lambda &= 0.68 \\ |\Omega_K| &\equiv \left| \frac{K}{H_0^2} \right| \leq 1 \times 10^{-2} \end{aligned}$$

So our universe is almost perfectly flat. From now on we will take this to be exact and set K=0. With this we can take (F1),

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \Omega a^{-3(1+\omega)} \tag{1.60}$$

and solve for a for each era:

$$a(t) \propto \begin{cases} t^{1/2} & \text{(RD)} \\ t^{2/3} & \text{(MD)} \\ e^{H_0 t} & (\Lambda\text{D}) \end{cases} \quad \text{or} \quad a(\tau) \propto \begin{cases} \tau & \text{(RD)} \\ \tau^2 & \text{(MD)} \\ -1/\tau & (\Lambda\text{D}) \end{cases} \tag{1.61}$$

Summary

In this chapter we have seen that making some very simple assumptions about the large scale universe, that it is homogeneous and isotropic, places tight constraints on both the curvature (metric) and constituents (EM tensor). We find that the metric must take the FRW form,

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-Kr^2} + r^2 d\Omega^2\right). \tag{FRW}$$

This led us to define the Hubble parameter,

$$H = \frac{\dot{a}}{a}. \tag{1.62}$$

With this metric the geodesic equation show us that momentum decays with expansion, which led us to define redshift.

$$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1} = \frac{1}{a} - 1 \tag{1.63}$$

Homogeneity and isotropy force the EM tensor to be that for a perfect fluid

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu + P g_{\mu\nu} \tag{1.64}$$

and it is described entirely by its equation of state $\omega = P/\rho$. Three equations of state were considered: Matter $\omega = 0$ (everything that moves slowly), Radiation $\omega = 1/3$ (everything that moves fast), and Dark Energy $\omega = -1$ (???).

Putting the metric and EM tensor together in Einsteins equations give the Friedmann equations

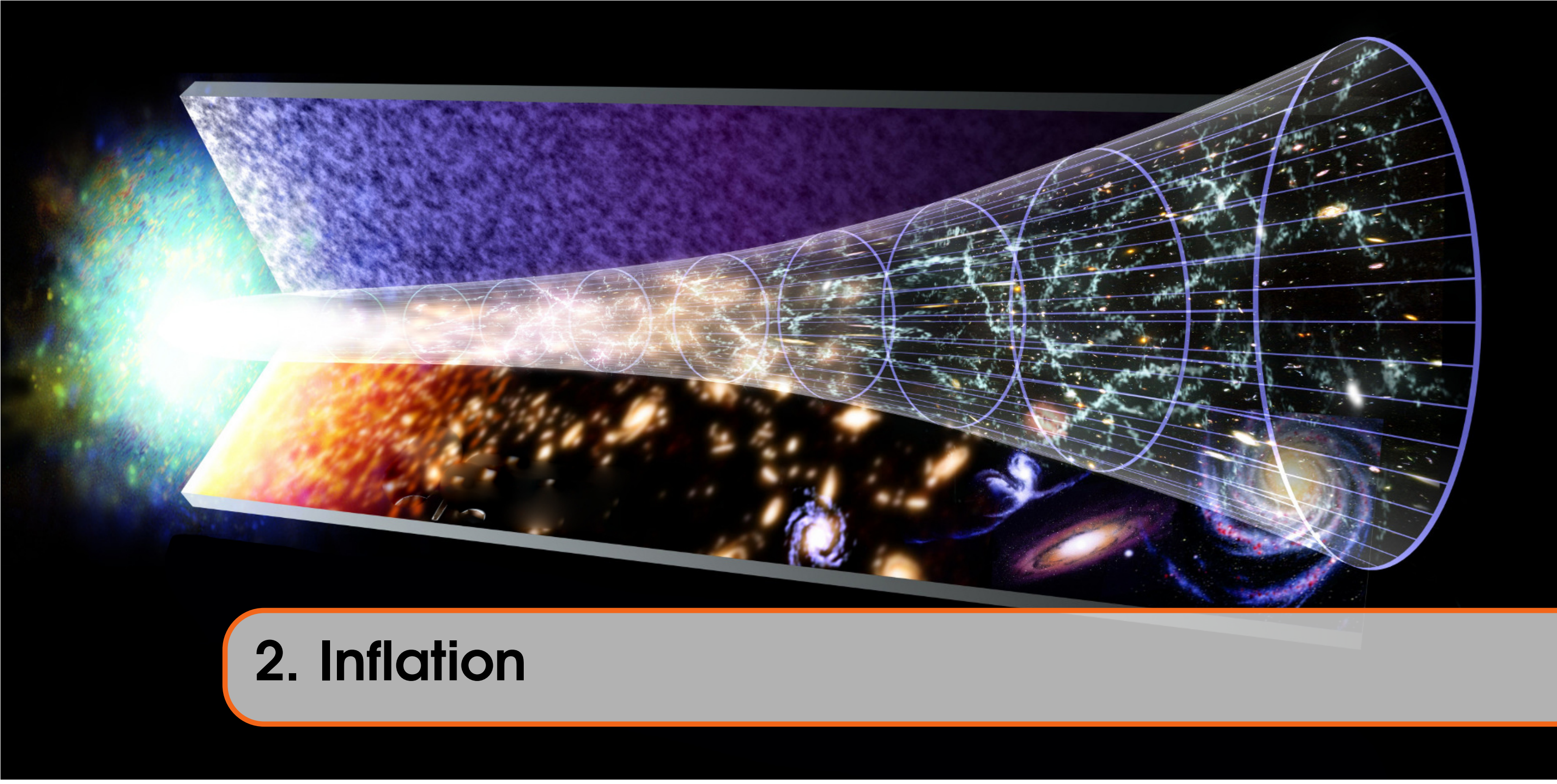
$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{F1}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{F2}$$

$$(\dot{F}1) + (F2) \implies \dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) \tag{F3}$$

Which describe the evolution of the scale factor a. Analysis of these equations shows that there is only one static solution, the ESU, which is unstable. So our universe must have begun with a “Big Bang" at time $t = 0$. When the universe is dominated by a single component we have the following solutions for the scale factor

$$a(t) \propto \begin{cases} t^{1/2} & \text{(RD)} \\ t^{2/3} & \text{(MD)} \\ e^{H_0 t} & (\Lambda\text{D}) \end{cases} \quad \text{or} \quad a(\tau) \propto \begin{cases} \tau & \text{(RD)} \\ \tau^2 & \text{(MD)} \\ -1/\tau & (\Lambda\text{D}) \end{cases} \tag{1.65}$$



2. Inflation

The Big Bang model has some notable successes:

1. **Big Bang Nucleosynthesis (BBN)**. If we we assume a thermal bath of standard model particles with a temperature $T > 100$GeV at early times then BBN correctly predicts the light element abundances.
2. **Cosmic Microwave Background (CMB)**. At early times the universe is hot so electrons are free and couple to photons via Thompson scattering. As the universe expands and cools the atoms combine to form neutral hydrogen and the photons free stream to today and are observed as the CMB

But it has many severe problems with initial conditions

1. **Flatness problem**. During matter and radiation dominated eras deviations from flatness grow with time (see examples sheets). So the fact that we observe the universe to be flat today requires extreme fine tuning of the early universe.
2. **Relic particles**. Most Grand Unified Theories (GUT) predict the production of “relic particles" like topological defects from phase transitions, like magnetic monoploes. Magnetic monopoles are very heavy and if produced would cause the universe to have collapsed.
3. **Horizon problem**. We observe the CMB to have an intensity which is uniform to 1 part in 10^5 yet it is composed of $30,000$ causally disconnected regions.
4. **.....**

Let us examine the horizon problem in detail.

2.1 Horizon Problem

First we need to define the particle horizon. If we take the FRW metric in conformal time and only consider radial motion (so dropping the $d\Omega$ part) we have

$$ds^2 = a^2(\tau)\left(-d\tau^2 + d\chi^2\right) . \tag{2.1}$$

Photons travel on null geodesics, $ds^2 = 0$, so we have $\Delta\chi = \pm\Delta\tau$. This implies that in conformal space photons travel in straight lines, see Figure 2.1 .

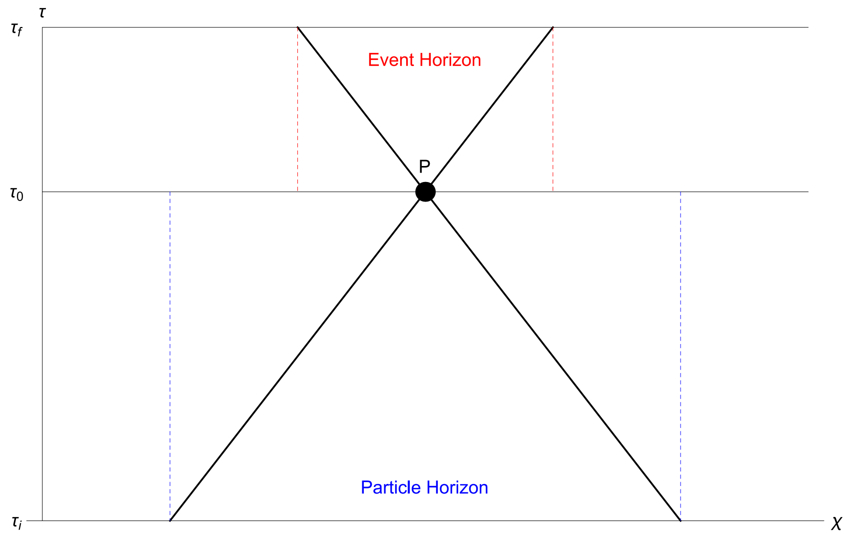


Figure 2.1: Particle and Event horizons for particle P

We can use this to calculate the physical particle horizon r_{PH},

$$r_{PH} = a\chi_{PH} = a(\tau_0 - \tau_i) = a\int_{t_i}^{t_0} \frac{1}{a} dt = a\int_{a_i}^{a_0} \frac{da}{a\dot{a}} = a\int_{\ln(a_i)}^{\ln(a_0)} \frac{1}{aH} d\ln(a) \tag{2.2}$$

In an expanding universe particles are moving apart via expansion so we are interested in the comoving particle horizon χ_{PH} as this determines if particles are moving into ($\dot{\chi}_{PH} > 0$) or out of ($\dot{\chi}_{PH} < 0$) the particle horizon as time passes

$$\chi_{PH} = \int_{\ln(a_i)}^{\ln(a_0)} \frac{1}{\mathscr{H}} d\ln(a) \tag{2.3}$$

where we have defined the comoving Hubble parameter $\mathscr{H} = aH = \frac{a'}{a}$ and $\mathscr{H}^{-1}$ is the comoving Hubble radius

---

Aside: We will use the following notation from now on for derivatives

$$\cdot = \frac{d}{dt} \qquad ' = \frac{d}{d\tau} \tag{2.4}$$

For a perfect fluid $a \propto t^{2/3(1+\omega)}$ so $\mathscr{H} = aH = \dot{a} = H_0 a^{-(1+3\omega)/2}$ and

$$\chi_{PH} = \frac{2H_0^{-1}}{1+3\omega}\left(a^{\frac{1+3\omega}{2}} - a_i^{\frac{1+3\omega}{2}}\right) = \frac{2}{1+3\omega}\mathscr{H}^{-1} \tag{2.5}$$

this leads to the confusing practice of calling both the particle horizon and the Hubble radius the “Horizon" as for ordinary matter they are related by a $O(1)$ constant. In general they are quite different. The **particle horizon**, which is everything we could have talked to, the **event horizon** which is everything we will be able to talk to in the future and the **Hubble radius** (H^{-1}) which is loosely the stuff we can communicate with “now". The easiest way to understand the Hubble radius is that it is the distance at which objects appear to move away at the speed of light, ie the apparent velocity v is:

$$v = \frac{dr}{dt} = \frac{d(a\chi)}{dt} = \dot{a}\chi = Hr \tag{2.6}$$

so $r = H^{-1}$ gives $v = 1$ $(= c)$ as we work in natural units. So we can see that objects further away than this will be difficult to communicate with. When we look at perturbations in the next section it will be very important whether their wavenumber is larger or smaller than H. In this course when we say "Horizon" we will invariably mean the comoving Hubble radius.

Now for matter domination we have $\chi_{PH} \propto \sqrt{a}$ so the comoving horizon grows with time. So the longer we wait the more we can see. This means that as time passes regions which are out of causal contact with each other will become visible to us.

Consider the CMB which we observe at $z \approx 1100$. By assuming matter domination back the big bang we can see that $t = 0 \implies \tau = 0$. Now we can draw Figure 2.2 and we see that the CMB consists of many causally disconnected regions. How many? The number of disconnected regions goes as the volume so we have

$$N_{disconnect} = \left(\frac{r_0}{r_{CMB}}\right)^3 = \left(\frac{a_0}{a_{CMB}}\right)^{\frac{3}{2}} = (1+z)^{\frac{3}{2}} \approx 30,000\,. \tag{2.7}$$

But we observe the CMB to be uniform to one part in 10^5.

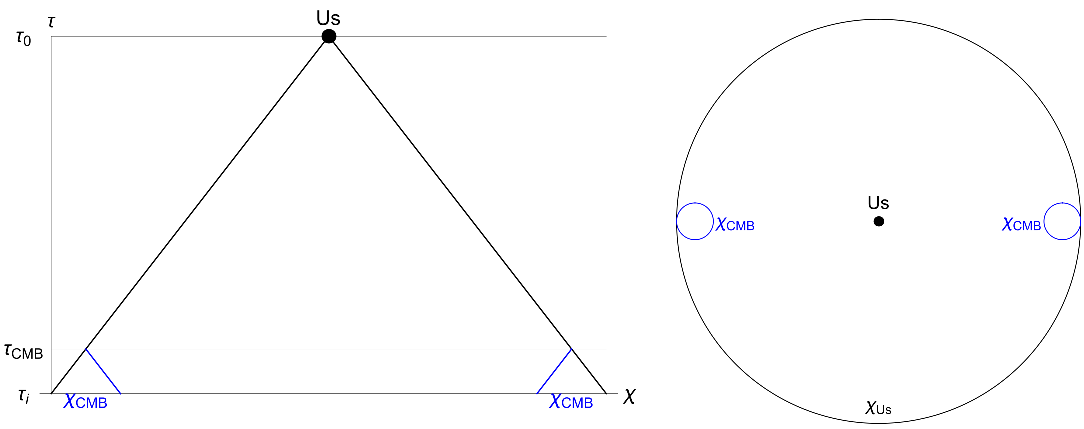


Figure 2.2: Particle Horizons for the CMB from the "side" on the left, and from "above" on the right. The horizon problem is that we could fit 30,000 blue circles into the black one on the right but observe the CMB to have an almost perfectly uniform temperature

2.2 Horizon Solution

The big bang model is very successful from BBN on so we only want to modify the evolution of the universe before t_{BBN}. The horizon problem is simply that the comoving horizon grows with time

$$\frac{d}{dt}\left(\mathscr{H}^{-1}\right) > 0 \tag{2.8}$$

as long as this happens we will eventually be able to see regions which are causally disconnected which we would normally expect to break homogeneity and isotropy. In fact, the horizon problem can be recast as in a universe filled with ordinary matter our twin assumptions of homogeneity and isotropy are very unnatural.

The solution to the horizon problem is simple. We need to have a period of time in the early universe where

$$\frac{d}{dt}\left(\mathscr{H}^{-1}\right) < 0 \tag{2.9}$$

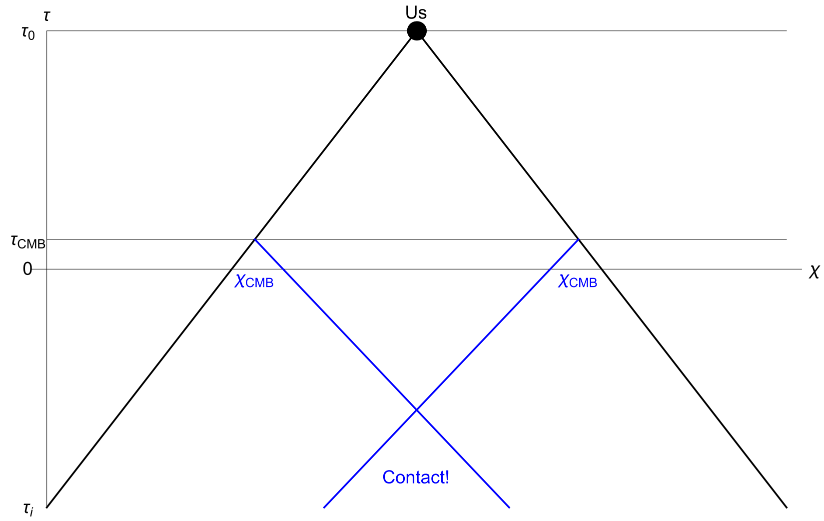


Figure 2.3: Now as $\tau_i \to -\infty$ we have enough conformal time for all regions of the CMB to have been in causal contact.

Then the comoving horizon is shrinking and casually connected regions exit the horizon. Then, when they re-enter at late times when the comoving horizon is growing, we don't need to be surprised that they look the same.

Now $\frac{d}{dt}\left(\mathscr{H}^{-1}\right) < 0 \implies (1+3\omega) < 0$ so we have that

$$\tau_i = \frac{2H_0^{-1}}{1+3\omega} a_i^{(1+3\omega)/2} \implies \tau_i \to -\infty \quad \text{as} \quad a_i \to 0 \tag{2.10}$$

so if we redraw our horizon plot for the CMB we see that all regions have previously been in contact, see Figure 2.3

The next question is how much do we need the comoving horizon to have shrunk? The answer is to explain the observed homogeneity and isotropy we observe we need all observable scales to have been in causal contact at the beginning of inflation. So we need

$$\mathscr{H}_i^{-1} > \mathscr{H}_0^{-1} \approx e^{60} \times \mathscr{H}_{BBN}^{-1} \tag{2.11}$$

and as $H \approx$ constant during inflation with $\mathscr{H} = aH$ we know that we need

$$a(\tau_{BBN}) > e^{60} a(\tau_i) \implies 60 \text{ e-folds of inflation} \tag{2.12}$$

Exercise: Inflation is often described a multitude of other other ways, show that the following are equivalent

1. **Period of Acceleration** $\ddot{a} > 0$

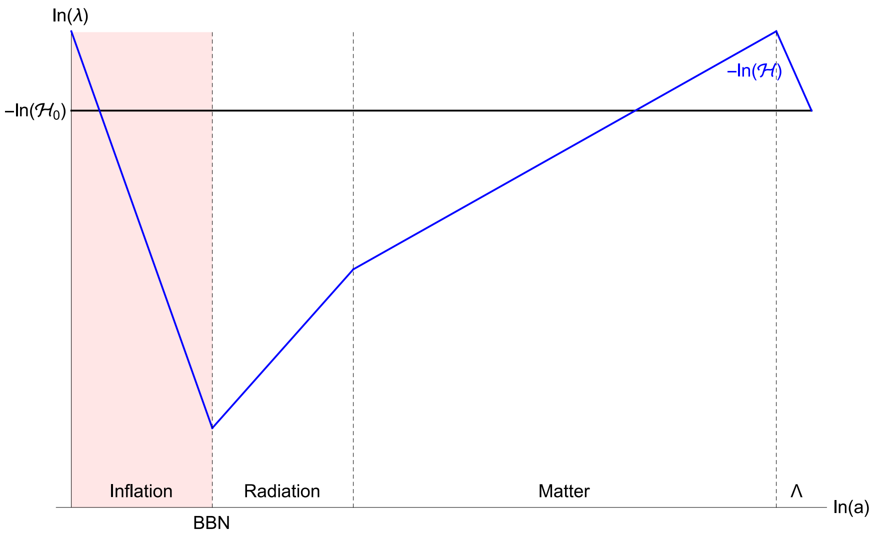


Figure 2.4: Plot of the evolution of the comoving Hubble radius with a for some comoving length scale λ. In order to have a Homogeneous and Isotropic universe today we need inflation (shaded red) to have lasted long enough that all observable scales ($\mathscr{H}_0^{-1}$) were in causal contact at the beginning of inflation, ie: $\mathscr{H}_i^{-1} > \mathscr{H}_0^{-1}$

2. **Slowly varying Hubble** $\varepsilon \equiv -\dot{H}/H^2 < 1$
3. **Exponential expansion** $a(t) \approx \exp(Ht)$
4. **Negative Pressure** $\omega < -1/3$
5. **Constant Density** $|\frac{d\ln(\rho)}{d\ln(a)}| = 2\varepsilon < 1$

2.3 Physics of inflation

For inflation to solve the initial condition problems identified in the Big Bang Model it needs to do 4 things

1. **Occur** We need a mechanisim which causes the comoving horizon to shrink. This is parameterised by

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} = -\frac{d\ln(H)}{d\ln(a)} < 1 \tag{2.13}$$

2. **Last** Once we have the conditions for inflation to begin we need it to last for at least 60-efolds. This is parameterised by

$$\eta \equiv \frac{\dot{\varepsilon}}{H\varepsilon} = \frac{d\ln(\varepsilon)}{d\ln(a)} < 1 \tag{2.14}$$

3. **End** Once inflation has lasted for 60-efolds needed to solve the horizon problem we need a natural mechanisim for inflation to end. This is called the “gracefull exit problem""
4. **Decay** Once inflation has ended the universe is empty and cold so we need the inflaton to decay and produce the thermal bath of standard model particles with temperature $T > 100\text{GeV}$ needed for BBN to begin. This process is called reheating.

The simplest models of inflation use a single scalar field ϕ (the inflaton) with energy density $V(\phi)$ (the inflaton potential).

The basic idea is that if we take ϕ to be roughly constant then as $V(\phi)$ does not dilute with the expansion of the universe if plays the roll of a constant energy density which we saw in the previous section causes exponential expansion. For a scalar field we have the following energy-momentum tensor

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \left(\frac{1}{2} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi + V(\phi) \right) \tag{2.15}$$

And for a homogeneous field $\phi = \phi(t)$ we can calculate

$$\rho_\phi \equiv -T^0_0 = \frac{1}{2}\dot{\phi}^2 + V(\phi) \quad \text{(KE + PE)} \tag{2.16}$$

$$P_\phi \equiv \frac{1}{3} T^i_i = \frac{1}{2}\dot{\phi}^2 - V(\phi) \quad \text{(KE - PE)} \tag{2.17}$$

We can substitute these into the Friedmann equations (using $M_{pl}^2 \equiv 1/8\pi G$) to get

$$(F1) \implies H^2 = \frac{1}{3M_{pl}^2} \left(\frac{1}{2}\dot{\phi}^2 + V(\phi) \right) \tag{2.18}$$

$$(F2) \implies \dot{H} = -\frac{1}{2M_{pl}^2} \dot{\phi}^2 \tag{2.19}$$

and taking $(\dot{F1}) + (F2)$ we obtain

Klein-Gordon Equation:

$$\underset{\text{(Acceleration)}}{\ddot{\phi}} + \underset{\text{(Friction)}}{3H\dot{\phi}} = \underset{\text{(Force)}}{-V_{,\phi}} \tag{KG}$$

where we have use the shorthand $V_{,\phi} = \frac{dV}{d\phi}$. Now the condition for inflation to occur is

$$\varepsilon = -\frac{\dot{H}}{H^2} = \frac{1}{M_{pl}^2} \frac{\dot{\phi}^2}{H^2} < 1 \tag{2.20}$$

so inflation occurs if the KE is small. This type of inflation is called “slow-roll" inflation. For inflation to last we need the acceleration to be small (so the KE remains small). If we define

$$\delta \equiv -\frac{\ddot{\phi}}{H\dot{\phi}} < 1 \tag{2.21}$$

then it is easy to show that

$$\eta = 2(\varepsilon - \delta) \tag{2.22}$$

so the conditions for inflation $(\varepsilon < 1\,, \eta < 1)$ and $(\varepsilon < 1\,, \delta < 1)$ are equivalent. This leads to the slow roll approximations, $\varepsilon \ll 1$ and $\delta \ll 1$, which allows us to neglect the kinetic term from (F1) and the acceleration term from (KG) giving

$$H^2 \approx \frac{V}{3M_{pl}^2} \tag{2.23}$$

$$3H\dot{\phi} \approx -V_{,\phi} \tag{2.24}$$

We can then use these to define the slow-roll parameters in terms of the potential

$$\varepsilon = \frac{M_{pl}^2}{2}\left(\frac{V_{,\phi}}{V}\right)^2 \equiv \varepsilon_V \tag{2.25}$$

$$2\varepsilon - \frac{1}{2}\eta = M_{pl}^2 \frac{V_{,\phi\phi}}{V} \equiv \eta_V \tag{2.26}$$

Aside: so we have three sets of slow-roll parameters ε, η, ε, δ, and $\varepsilon_V \eta_V$. The first set are general conditions for a shrinking Hubble sphere defining inflation, the second are an equivalent formulation and the last in terms of the potential are less general and only hold for slow roll.

And we can use them to calculate the total amount of inflation that occurs,

$$N_{tot} = \int_{a_i}^{a_f} d\ln(a) = \int_{t_i}^{t_f} H dt \approx \int_{\phi_i}^{\phi_f} \frac{1}{\sqrt{2\varepsilon_V}} \frac{|d\phi|}{M_{pl}} \geq 60 \tag{2.27}$$

We can see that the number of e-folds of inflation is proportional to $1/\sqrt{\varepsilon_V}$ so small ε_V means large N_{tot}. The effect of η_V is on the duration of inflation so the smaller η_V the further in further apart ϕ_i and ϕ_f are leading to larger N_{tot}.

Now the slow-roll parameters are defined purely in terms of the shape of the potential. So what type of potentials do we need?

There are 100+ models of inflation but one simple way to classify them is into “small field" and “large field", see Figure 2.5.

1. **Small field** (or plateau) models typically we have $V_{,\phi\phi} < 0$ with potentials like $V \approx 1 - (\frac{\phi}{\mu})^n$ (think Taylor expansion around the origin) where μ is some mass scale and n a small power. Here we force the slow roll parameters, ε_V and η_V to be small by choosing a very flat potential so $V_{,\phi}$ and $V_{,\phi\phi}$ are small.
2. **Large field** models we typically have $V_{,\phi\phi} > 0$ with potentials like $V \approx \phi^n$ or $V \approx e^{\frac{\phi}{\mu}}$. Here we force the slow roll parameters, ε_V and η_V to be small by making the potential large, so $1/V$ is small.

The reason these two classifications are useful is because they are distinguishable via there predictions for the production of gravitational radiation. Gravity waves are produced during inflation with an amplitude which is directly related to the energy scale of inflation ($\propto V^{\frac{1}{4}}$). In large field models we must start high up the potential so at high energy and so the amplitude of gravity waves is large. In small field models we have inflation at quite low energies so the amplitude of gravity waves is small. The amplitude of gravity waves is parameterised by r, the tensor to scalar ratio.

Now that we have made inflation happen we need to find a way for it to end and connect up with the standard big bang model. Inflation will end when the kinetic energy exceeds the potential energy for the inflaton. This occurs in the small field models when the potential gets too steep, or in large field models when we have dropped too low. Finally the inflaton reaches a minimum of the potential where it is trapped. We can always expand the potential around the minimum and approximate it with,

$$V_{min} \approx \frac{1}{2}m^2\phi^2 \tag{2.28}$$

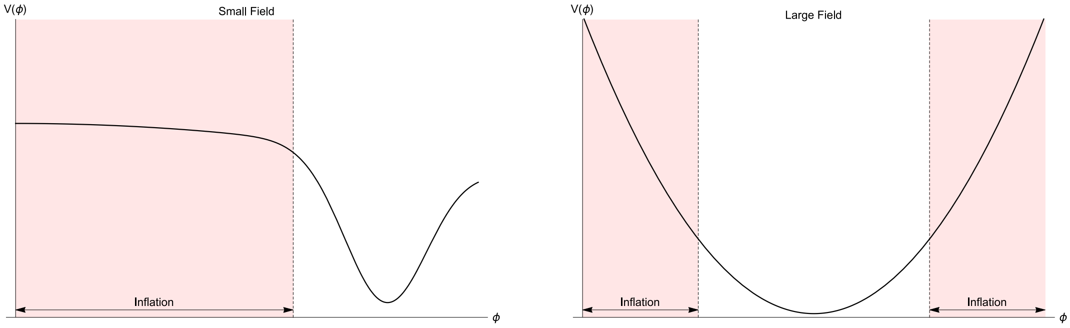


Figure 2.5: Two possible potential types for inflation, small field on the left and large field on the right

we can substitute this into the (KG) equation to get

$$\ddot{\phi} + 3H\dot{\phi} + m^2\phi = 0 \tag{2.29}$$

Now inflation has ended we can neglect the friction term and we find the solution is $\phi \approx \phi_0 \cos(mt)$ so the field oscillates. We can calculate $\langle \rho_\phi \rangle$ and $\langle \dot{\rho}_\phi \rangle$

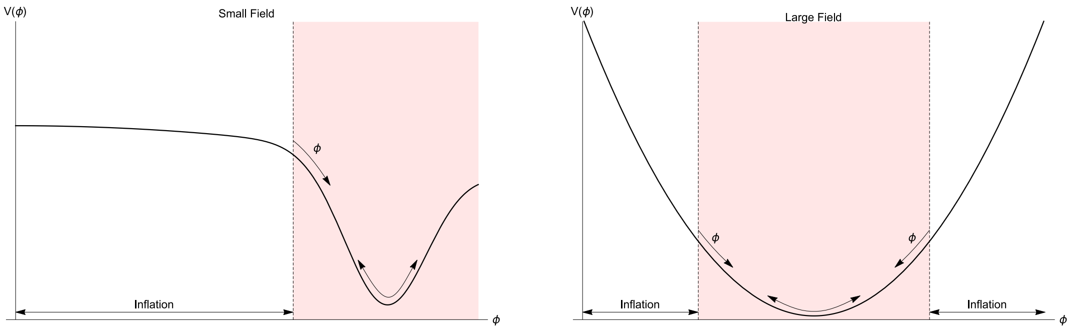


Figure 2.6: How inflation ends in the two example cases. The field exits the inflationary region and the universe stops inflating then the field reached a minimum of the potential where it oscillates.

$$\langle \rho_\phi \rangle = \frac{1}{2}\langle \dot{\phi}^2 + m^2\phi^2 \rangle \approx \langle \dot{\phi}^2 \rangle \quad \text{(equipartition of energy)} \tag{2.30}$$

$$\langle \dot{\rho}_\phi \rangle = \langle \dot{\phi}\ddot{\phi} + V_{,\phi}\dot{\phi} \rangle = \langle \dot{\phi}\left(\ddot{\phi} + V_{,\phi}\right) \rangle = -3H\langle \dot{\phi}^2 \rangle \tag{2.31}$$

$$\implies \langle \dot{\rho}_0 \rangle = -3H\langle \rho_\phi \rangle \tag{2.32}$$

and so $\langle \rho_\phi \rangle \propto a^{-3}$ and the oscillations behave like matter. The universe is now large, cold, flat and empty apart from a single oscillating scalar field. We need to find a way to transfer the energy in the inflaton to the standard model particles so BBN can begin.

2.4 Reheating*

In order for us to transfer the energy stored in the inflaton to SM fields we need to introduce coupling between them (or to some dark sector fields which then decay into SM particles). To begin with we will take a perturbative treatment. Assume we have some interaction Lagrangian $\mathscr{L}_{int}$ which has an associated decay rate Γ into particles χ. This modifies the (KG) equation to include an extra friction term due to decay,

$$\ddot{\phi}+(3H+\Gamma)\,\dot{\phi}=-V_{,\phi} \tag{2.33}$$

now when inflation ends $3H\leq\Gamma$, and particle production becomes significant. The eventually the inflation decays completely and we have a universe with reheat temperature $T_R\propto\sqrt{\Gamma M_{pl}}$. Unfortunately this perturbative approach fails in two ways.

1. It is too slow so T_R is typically below the 100GeV we need for BBN
2. Neither ϕ nor χ is in a thermal superposition of states needed for the perturbative treatment to apply. ϕ is a coherent oscillating field and χ is in the vacuum state so should be treated quantum mechanically

A better treatment is to take ϕ as a classical oscillating background with to which χ is coupled and is produced quantum mechanically. The modes functions then satisfy an equation of the form

$$\ddot{\chi}_k+\left(k^2+m_\chi^2+g^2A_\phi^2\sin^2(m_\phi t)\right)\chi_k=0 \tag{2.34}$$

where m_χ is the mass of χ, m_ϕ is the mass of ϕ, g is the coupling constant, A_ϕ is the amplitude of the oscillation of ϕ. This can be re-written as a Mathieu equation,

$$\chi_k''+(A_k-2q\cos(2z))\,\chi_k=0 \tag{2.35}$$

where $z=m_\phi t$, $A_k=2q+(k^2+m_\chi^2)/m_\phi^2$, and $q=g^2A_\phi^2/4m_\phi$. The Mathieu equation describes parametric resonance in forced oscillators and so for some values of k we expect there to be resonances which lead to exponential growth of χ_k. This process can be very efficient and results in rapid energy transfer leading to a reheat temperature $T_R\gg 100$GeV. This is called **preheating**.

As the process is very rapid the χ_k are produced out of thermal equilibrium and we then need a period of thermalisation (and decay to SM particles is χ is a dark sector field). Provided there is sufficient time for this to occur we can then begin BBN and in general there are few observational consequences which is why we have not looked at it in more detail. When there are some they generally come from the non-adiabatic energy transfer which can produce gravity waves or induce non-Gaussianities in the density perturbation amongst others but these tend to be fairly model dependent. For a review of this area see ArXiv:1001.2600.

2.5 Issues with inflation?

We have shown that inflation can solve the initial condition problems of the big bang model:

1. **Flatness** exponential expansion dilutes curvature
2. **Relics** exponential expansion dilutes pre-existing particle number densities to zero
3. **Horizon** Observable universe is inside the horizon at the beginning of inflation

But what about the initial conditions for inflation? Do we have any problems with them? This is a much debated area, some of the usual critisims are

1. What is ϕ? There are lots of candidates, compactification in string theory produces large numbers of scalar fields. The real question is why is only one dynamical while the others are held fixed? And is this configuration stable to, for example, de-compactification?
2. Is V natural? For inflation to work we need a potential which is very smooth and relatively flat. However 60 e-folds of inflation usually involves moving more than a Planck length in field space so we should typically expect corrections to the potential which would ruin this.
3. Given a theory, how likely is inflation? Of all the possible positions we could begin how likely is it that we start with a ϕ and $\dot{\phi}$ which allow inflation? This leads into the “Measure Problem" which is about what is the right way to assign probability to initial conditions.

The real answer (or at least my one) is that we don't know. Without a UV complete theory and some knowledge of what a typical realisation should look like from which to derive inflation we can only make guesses. Also there is the question of how likely do we need inflation to be? We only know that it happened once and whatever the probability you can always make anthropic arguments as we wouldn't exist in a universe without it.

So should we really believe inflation? Well there is one other clear, natural prediction of inflation which is the production of a scale invariant spectrum of density perturbations from which all structure we observe today grew. We will examine this process in detail in chapter 6 but I will sketch it here as we will need the power spectrum for the next two chapters.

During inflation the field ϕ is slowly rolling down the potential V thus the field acts as a global clock. However global clocks aren't allowed in quantum field theories so we expect the field to have small quantum fluctuations from location to location. These lead to small density perturbations as the quantum perturbation change how long inflation lasts for, eg:

$$\begin{aligned}\phi \text{ fluctuates up the potential} &\implies \text{Inflation is longer} \implies \text{lower density}\\ \phi \text{ fluctuates down the potential} &\implies \text{Inflation is shorter} \implies \text{higher density}\end{aligned}$$

we can make a reasonable guess at the spectrum of the fluctuations with a simple argument. If we divide the inflationary region in field space into regions of equal size $\Delta\phi$, see Figure 2.7

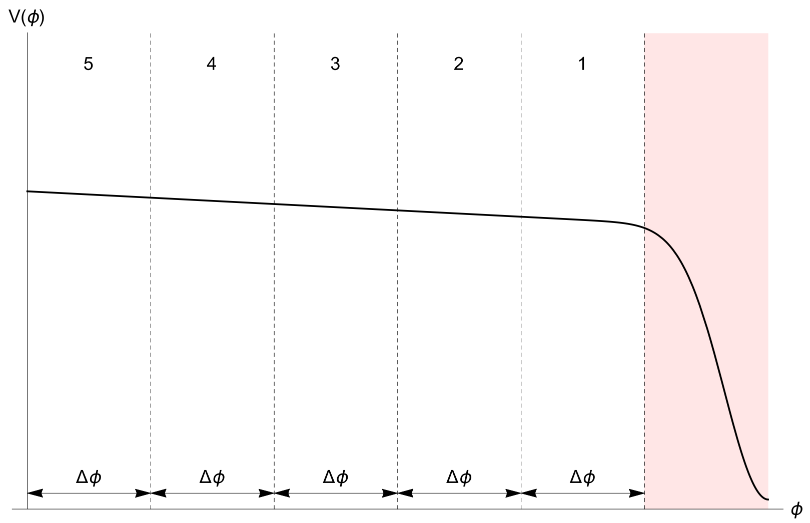


Figure 2.7: For the small field case we divide the inflationary region into five parts of equal size $\Delta\phi$

Now in the slow roll approximation we have $\ddot{\phi}$ to be very small so we expect $\dot{\phi} \approx$ constant so we traverse regions of equal ϕ in equal time so each region has equal Δt. Also, the universe is undergoing exponential expansion, $a \propto e^{Ht}$ with $H \approx$ constant, so in each region the universe undergoes equal log expansion $\Delta\ln(a)$, or equivalently in each region the universe grows by the same multiplicative factor $A = e^{H\Delta t}$).

As each region is effectively identical we expect the quantum fluctuations generated in each to have the same characteristic size, say λ, and amplitude. Now these fluctuations are stretched by the rapid expansion beyond the horizon size where they become classical. Considering Figure 2.7 again fluctuations produced in region 5 traverse 5 regions so grow to size $\lambda_5 \approx A^5\lambda$, those in region 4 traverse 4 regions grow to size $\lambda_5 \approx A^4\lambda$, region 3 $\lambda_3 \approx A^3\lambda$, region2 $\lambda_2 \approx A^2\lambda$, and region 1 $\lambda_1 \approx A^1\lambda$. So each region corresponds to an equal log interval of scales. We expect the power spectrum of quantum perturbations created in each to be the same so at the end of inflation we expect to have produced a spectrum of density perturbations with equal power in equal log intervals which is known as a scale invariant spectrum. To summarise:

$$
\begin{aligned}
(\Delta\phi)_i = (\Delta\phi)_j &\implies (\Delta t)_i = (\Delta t)_j \\
&\implies (\Delta \ln a)_i = (\Delta \ln a)_j \\
&\implies (\Delta \ln \lambda)_i = (\Delta \ln \lambda)_j \\
&\text{so} \\
\text{“Equal intervals of } \phi\text{"} &\implies \text{“Equal log intervals of scale"} \\
&+ \\
V(\phi) \approx \text{const.} &\implies \text{“Equal power in each interval"} \\
&= \\
&\text{“Scale invariant spectrum"}
\end{aligned}
$$

What does a scale invariant spectrum look like? If we define the power spectrum of inflaton fluctuations, $P_{\delta\phi}(k)$, via

$$\left\langle \delta\phi(\mathbf{k})\delta\phi^*(\mathbf{k}')\right\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}')P_{\delta\phi}(k) \tag{2.36}$$

where k is the wavenumber of the fluctuation. The power in a particular interval is

$$\int_{k_i}^{k_f} P_{\delta\phi}(k)d^3k = 4\pi\int_{k_i}^{k_f} P_{\delta\phi}(k)k^2dk \tag{2.37}$$

$$= 4\pi\int_{\ln k_i}^{\ln k_f} \left[k^3P_{\delta\phi}(k)\right] d\ln k \tag{2.38}$$

so for perfect scale invariance we need $P_{\delta\phi}(k) \propto k^{-3}$.

In reality V isn't constant but decreasing slowly (otherwise we wouldn't move) so as (F1) gives us $H^2 \propto V$ and the expansion is $a \propto e^{Ht}$ we have less expansion in later regions. This means that we lose some small scale power (of order ε). We can parameterise this by defining a spectral index n_s by

$$P^{\delta\phi}(k) \propto k^{n_s-4} \tag{2.39}$$

where $n_s = 1$ is scale invariance, and we expect (n_s-1) to be $O(\varepsilon)$ and < 0.

Summary

In this section we saw that the big bang model has severe problems with initial conditions. These are known as:

- Flatness Problem
- Relic Problem
- Horizon Problem

We examined the horizon problem in detail defining both the comoving particle horizon χ_{PH} and the comoving Hubble radius $\mathscr{H}^{-1}$ and saw that

$$\chi_{PH} = \int \mathscr{H}^{-1} d\ln(a)\,, \quad \mathscr{H} = \dot{a} \tag{2.40}$$

The comoving Hubble radius always grows and for a universe filled with ordinary matter $\omega > -1/3$,

$$\frac{d}{dt}\left(\mathscr{H}^{-1}\right) > 0\,. \tag{2.41}$$

This implies that as time goes on we can see particles which can never have communicated.

We saw that the solution is to propose a period in the early universe where the comoving horizon shrinks,

$$\frac{d}{dt}\left(\mathscr{H}^{-1}\right) < 0 \tag{2.42}$$

which will happen if $\omega < -1/3$ and is called **INFLATION**. If we require that the comoving Hubble radius shrinks as much during inflation as it has grown in the big bang phase then all regions we observe would originally have been in causal contact. We found that this required 60 e-folds of inflation.

Next we defined two parameters which tell us if inflation is successful

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} = -\frac{d\ln(H)}{d\ln(a)} < 1 \tag{2.43}$$

$$\eta \equiv \frac{\dot{\varepsilon}}{H\varepsilon} = \frac{d\ln(\varepsilon)}{d\ln(a)} < 1 \tag{2.44}$$

ε tells us if inflation occurs and η tells us if inflation lasts. We saw that inflation can be driven by a scalar field and used the Friedmann equations to derive the Klein Gordon equation

$$\ddot{\phi} + 3H\dot{\phi} = -V_{,\phi} \tag{KG}$$

and saw that $\varepsilon < 1$ required that the kinetic energy is small. If we make the slow roll approximation ($\varepsilon \ll 1\,, \eta \ll 1$) then we could write down the slow roll equations

$$H^2 = \frac{V}{3M_{pl}} \tag{2.45}$$

$$3H\dot{\phi} = -V_{,\phi} \tag{2.46}$$

and we could define the slow roll parameters in terms of the potential

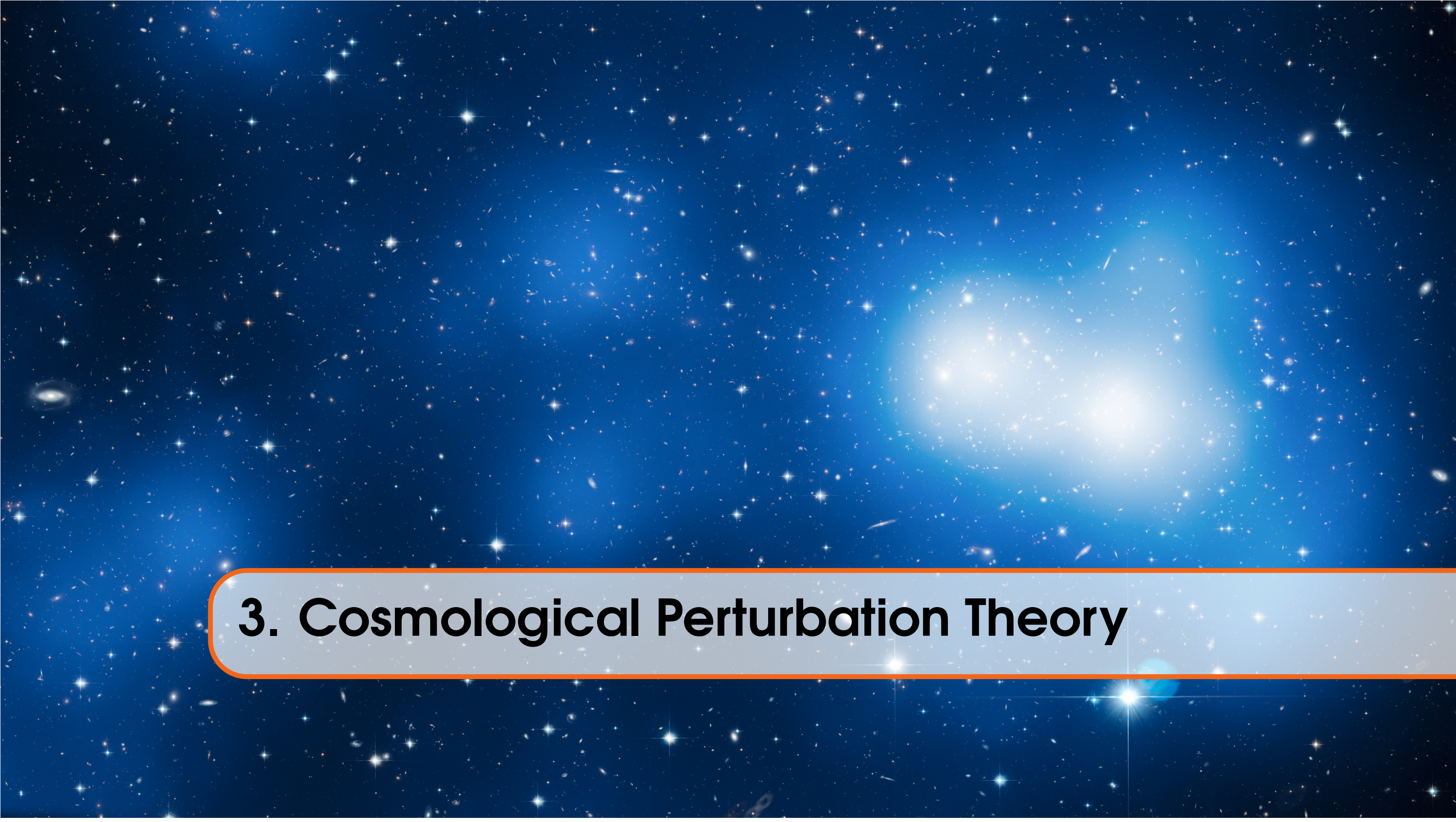
$$\varepsilon_V = \frac{M_{pl}^2}{2}\left(\frac{V_{,\phi}}{V}\right)^2 \tag{2.47}$$

$$\eta_V = M_{pl}^2 \frac{V_{,\phi\phi}}{V} \tag{2.48}$$

We discussed two types of potentials, small-field and large-field. The first makes ε, η small by making the derivatives small. The second makes ε, η small by making the potential large. These two classes of models have one key observational difference which is that the large-field models produce much more gravitational radiation than the small-field ones.

In both, inflation ends when the field rolls into a local minimum and oscillates with an energy density that behaves like matter. We discussed how the oscillations must decay into a thermal bath of SM particles to allow BBN to occur via a process called Reheating.

Finally we discussed possible issues with the initial conditions for inflation and made a heuristic argument that they produce a scale invariant spectrum of density perturbations.



3. Cosmological Perturbation Theory

We wish to derive the equations governing the evolution of perturbations to the simple background we explored in Chapter 1. We will do this by dividing all quantities, eg X, into a background part $\bar{X}$ and a perturbation δX. Having perturbed all quantities we will then substitute them into the Einstein equations and split out the background parts (which will be the Friedmann equations we met earlier) and the perturbation equations. Let us begin with the metric.

3.1 Perturbed metric

We can write the most general perturbation of the FRW metric as

$$ds^2 = a^2(\tau)\left(-(1+2A)\,d\tau^2 + 2B_i dx^i d\tau + (\delta_{ij}+h_{ij})\,dx^i dx^j\right) \tag{3.1}$$

Note we don't need to perturb a as we can absorb it into the time perturbation also all perturbations are function of time and space, $A = A(\tau, \mathbf{x})$ etc...

We can make a Scalar-Vector-Tensor (SVT) decomposition. This is particularly useful as at linear order they decouple in the Einstein equations and evolve separately. To do this we define

$$A = A \tag{3.2}$$

$$B_i = \partial_i B + B_i^V \tag{3.3}$$

$$h_{ij} = 2C\delta_{ij} + 2\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E + \left(\partial_i E_j^V + \partial_j E_i^V\right) + 2E_{ij}^T \tag{3.4}$$

where vector quantities are divergence free and tensor quantities are transverse and traceless so

$$\partial^i B_i^V = \partial^i E_i^V = 0 \tag{3.5}$$

$$\partial^i E_{ij}^T = \delta^{ij} E_{ij}^T = 0 \tag{3.6}$$

and we have split the scalar part of the tensor $h_{ij} = \partial_i\partial_j H$ into a trace part, C, and trace free part, E. We know we have 10 degrees of freedom in the metric (as it's 4x4 and symmetric) so let's count and see if we have them all.

$$\begin{aligned} 4\,\text{scalars}\,(A,B,C,E)\,\text{with 1 dof each} &= 4\,\text{dof} \\ 2\,\text{vectors}\,(B_i^V,E_i^V)\,\text{with 2 dof each} &= 4\,\text{dof} \\ 1\,\text{tensor}\,(E_{ij}^T)\,\text{with 2 dof each} &= 2\,\text{dof} \end{aligned}$$

and $4+4+2=10$ so we have accounted for all possible perturbations (which also justifies us not perturbing a).

Aside: The SVT decomposition can also be interpreted as a spin decomposition where we have 4 spin-0 fields (scalars), 4 spin-1 fields (vectors) and 2 spin-2 fields (tensors). It's a bit fiddly to show but not too difficult. You need to work in Fourier space then consider how they behave under spatial rotations. Anthony Challinor has most if it in his Part III cosmology notes from 2009 (Google it). You should also note that this decomposition is unique (which is why they end up separating in the Einstein equations).

Now the vectors describe rotational velocity modes or "Vorticity". These are not sourced during inflation and even if they were they always decay with expansion as a^{-2} (it's just conservation of angular momentum, if you increase the radius by a factor A you rotate A^{-2} times as fast). Tensors perturbations are gravity waves with the two dof being the two polarisation states. We will touch on them in chapter 6 and on your examples sheets. Here we will neglect both and focus solely on scalar perturbations.

Just considering scalars we have the perturbed metric taking the form

$$ds^2 = a^2(\tau)\left(-(1+2A)\,d\tau^2 + 2\partial_i B dx^i d\tau + \left[(1+2C)\,\delta_{ij} + 2\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E\right]dx^i dx^j\right) \tag{3.7}$$

Unfortunately that's not the end of it, two of these perturbations will turn out to be "fake" as we will see in the next section

3.2 Gauge Problem

Now let us consider an unperturbed universe with metric

$$ds^2 = a^2(\tau)\left(-d\tau^2 + \delta_{ij}dx^i dx^j\right) \tag{3.8}$$

If we make a change of coordinates

$$x^i \to \tilde{x}^i = x^i + \xi^i \implies dx^i = d\tilde{x}^i - \partial_\tau \xi^i d\tau - \partial_k \xi^i dx^k \tag{3.9}$$

then the metric takes the form

$$ds^2 = a^2(\tau)\left(-d\tau^2 - 2\xi_i' d\tilde{x}^i d\tau + \left(\delta_{ij} - 2\partial_{(i}\xi_{j)}\right)d\tilde{x}^i d\tilde{x}^j\right) \tag{3.10}$$

and we have induced fake perturbations ξ_i' (scalar) and $\partial_{(i}\xi_{j)}$ (vector). These are known as **gauge modes**. Similarly we can perturb time to get

$$\tau = \tilde{\tau} + \xi^0 \implies \rho(\tilde{\tau}+\xi^0) = \rho(\tau) + \xi^0\rho'(\tau) \tag{3.11}$$

where we have induced a fake density perturbation.

Let us consider the general case of a gauge transformation

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu \quad \xi^0 \equiv T, \xi^i \equiv \partial^i L \tag{3.12}$$

where we have only considered scalar perturbations (In general we would also have a vector perturbation $\xi_i^V \equiv L_i^V$ but no tensor part. This allows us to immediately deduce that the tensor part is gauge invariant and not affected by this problem, the vector equations however will have one gauge mode) Now we can deduce the effect of the gauge transformation on the metric by remembering that

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = \tilde{g}_{\alpha\beta} d\tilde{x}^\alpha d\tilde{x}^\beta \tag{3.13}$$

so we have

$$g_{\mu\nu} = \frac{d\tilde{x}^\alpha}{dx^\mu} \frac{d\tilde{x}^\beta}{dx^\nu} \tilde{g}_{\alpha\beta} \tag{3.14}$$

Let us consider the 00 part of the metric

$$\begin{aligned} g_{00} &= \left(\frac{d\tilde{x}^0}{dx^0}\right)^2 \tilde{g}_{00} \\ a^2(1+2A) &= \left(1+T'\right)^2 a^2(\tau+T)\left(1+2\tilde{A}\right) \\ &\approx \left(1+2T'\right)(1+2\mathscr{H}T)a^2(\tau)\left(1+2\tilde{A}\right) \\ \tilde{A} &= A - T' - \mathscr{H}T \end{aligned} \tag{3.15}$$

Exercise: Show the following

$$\tilde{B} = B + T - L' \tag{3.16}$$

$$\tilde{C} = C - \mathscr{H}T - \frac{1}{3}\nabla^2 L \tag{3.17}$$

$$\tilde{E} = E - L \tag{3.18}$$

So how do we solve the gauge problem? There are a few options:

1. Compute Observables.
2. Follow everything, metric and matter and identify the gauge modes.
3. Fix the Gauge. There are three popular metric gauges.
   (a) **Synchronous**, $A = B = 0$, where time is unperturbed.
   (b) **Spatially Flat**, $C = E = 0$, where space is unperturbed.
   (c) **Longitudinal/Newtonian**, $B = E = 0$, which has no off diagonal part (hence Longitudinal) and the spatial perturbation is directly analogous to the potential in Newtonian gravity (hence Newtonian)
4. Work with Gauge invariant quantities, eg. the Bardeen potentials:

$$\Psi_B \equiv A + \mathscr{H}\left(B - E'\right) + \left(B - E'\right)' \tag{3.19}$$

$$\Phi_B \equiv -C - \mathscr{H}\left(B - E'\right) + \frac{1}{3}\nabla^2 E \tag{3.20}$$

Exercise: Show Ψ_B and Φ_B are gauge invariant

In Newtonian gauge we have $\Psi_B = A$ and $\Phi_B = -C$ and the metric takes the form

$$ds^2 = a^2\left(-(1+2\Psi)d\tau^2 + (1-2\Phi)\,\delta_{ij}dx^i dx^j\right) \tag{3.21}$$

where we have dropped the $_B$ for convenience.

3.3 Perturbed Matter

Now we need to perturb the matter. If we start with the background energy momentum tensor

$$\bar{T}^{\mu}_{\nu} = (\bar{\rho}+\bar{P})\,\bar{U}^{\mu}\bar{U}_{\nu} + \bar{P}\delta^{\mu}_{\nu} \tag{3.22}$$

we can write down the perturbed energy momentum tensor

$$\delta T^{\mu}_{\nu} = (\delta\rho + \delta P)\,\bar{U}^{\mu}\bar{U}_{\nu} + (\bar{\rho}+\bar{P})\,(\delta U^{\mu}\bar{U}_{\nu} + \bar{U}^{\mu}\delta U_{\nu}) + \delta P\delta^{\mu}_{\nu} + \Pi^{\mu}_{\nu} \tag{3.23}$$

Where we have added Π^{μ}_{ν} as a catch-all for any possible perturbation we might have missed. So what could we have missed? We can absorb any part of $\Pi_{\mu\nu}$ which is in the same direction as U^{μ} into the energy density so $U^{\mu}\Pi_{\mu\nu} = 0 \implies \Pi_{0\nu} = 0$ and we only have the Π_{ij} left. Furthermore we can also absorb the trace of Π_{ij} into the pressure perturbation so Π_{ij} must be traceless. This means that Π_{ij} is the anisotropic stress (pressure being the isotropic stress). If we only consider scalar quantities we can write $\Pi_{ij} = \left(\partial_i\partial_j - \frac{1}{3}\nabla^2\delta_{ij}\right)\Pi$.

Now we need to work out what δU^{μ} looks like. We have $\bar{U}^{\mu} = \frac{1}{a}(1,\mathbf{0})$ and we know that $g_{\mu\nu}U^{\mu}U^{\nu} = -1$ so we can perturb this to find

$$\begin{aligned}\delta g_{\mu\nu}\bar{U}^{\mu}\bar{U}^{\nu} + 2\bar{U}_{\mu}\delta U^{\mu} &= 0\\ -2A - 2a\delta U^0 &= 0\\ \delta U^0 &= -\frac{1}{a}A\end{aligned} \tag{3.24}$$

Now we are free to define δU^i so we will make the logical choice

$$U^{\mu} = \frac{1}{a}\left(1-A, \mathbf{v}^i\right) \tag{3.25}$$

where $\mathbf{v}^i \equiv \frac{dx^i}{d\tau}$ which, as we are only dealing with scalars, we can write as $\mathbf{v}_i = \partial_i v$

Exercise: Show the following

$$U_{\mu} = a\left(-(1+A), \partial_i(v+B)\right) \tag{3.26}$$

$$\delta T^0_0 = -\delta\rho \tag{3.27}$$

$$\delta T^0_i = (\bar{\rho}+\bar{P})\,\partial_i(v+B) \tag{3.28}$$

$$\delta T^i_0 = -(\bar{\rho}+\bar{P})\,\partial^i v \equiv -\partial^i q \quad \text{(the 3-momentum density)} \tag{3.29}$$

$$\delta T^i_j = \delta P\delta^i_j + \left(\partial^i\partial_j - \frac{1}{3}\nabla^2\delta^i_j\right)\Pi \tag{3.30}$$

In a multi-component universe the total energy momentum tensor is the sum of the individual energy momentum tensors so we can deduce

$$\delta\rho = \sum_I \delta\rho_I\,, \quad \delta P = \sum_I \delta P_I\,, \quad \delta q = \sum_I \delta q_I\,, \quad \delta\Pi = \sum_I \delta\Pi_I \tag{3.31}$$

Note we can only add the 3-momentum densities not the velocities.

We can check how all the perturbations transform under a general gauge transformation

$$T^\mu_\nu = \frac{dx^\mu}{d\tilde{x}^\alpha}\frac{d\tilde{x}^\beta}{dx^\nu}\tilde{T}^\alpha_\beta \tag{3.32}$$

Exercise: Show the following

$$\tilde{\delta\rho} = \delta\rho - T\bar{\rho}' \tag{3.33}$$

$$\tilde{\delta P} = \delta P - T\bar{P}' \tag{3.34}$$

$$\tilde{v} = v + L' \tag{3.35}$$

$$\tilde{\Pi} = \Pi \tag{3.36}$$

There are two popular matter gauges

1. **Uniform Density**, $\delta\rho = B = 0$, where spatial slices follow surfaces of constant density.
2. **Comoving**, $v = B = 0$, where spatial slices move with the fluid.

and one popular gauge invariant quantity,

$$\Delta \equiv \frac{\delta\rho}{\bar{\rho}} + \frac{\bar{\rho}'}{\bar{\rho}}(v+B) \tag{3.37}$$

which in comoving gauge is $\frac{\delta\rho}{\bar{\rho}} \equiv \delta$ where we have defined δ the density contrast. So Δ has the snappy name “the comoving gauge density contrast"

Now we have perturbed all quantities we can use substitute them into Einstein equation to obtain our perturbed equations.

3.4 Linearised Equations

We will now specialise to the Newtonian gauge (You will calculate the equations in there general form using the $3+1$ formalism in the Advanced Cosmology course) where the metric tensor has the simple form,

$$g_{\mu\nu} = a^2\begin{pmatrix} -(1+2\Psi) & 0 \\ 0 & (1-2\Phi)\delta_{ij}\end{pmatrix}, \quad g^{\mu\nu} = \frac{1}{a^2}\begin{pmatrix} -(1-2\Psi) & 0 \\ 0 & (1+2\Phi)\delta^{ij}\end{pmatrix} \tag{3.38}$$

We can easily (if labouriously) calculate the Christoffel symbols

Exercise: Show the following

$$\Gamma^0_{00} = \mathscr{H} + \Psi' \tag{3.39}$$
$$\Gamma^0_{0i} = \partial_i \Psi \tag{3.40}$$
$$\Gamma^0_{ij} = \left(\mathscr{H} - \left[\Phi' + 2\mathscr{H}(\Phi+\Psi)\right]\right)\delta_{ij} \tag{3.41}$$
$$\Gamma^i_{00} = \partial^i \Psi \tag{3.42}$$
$$\Gamma^i_{0j} = \left(\mathscr{H} - \Phi'\right)\delta^i_j \tag{3.43}$$
$$\Gamma^i_{jk} = \left(\delta_{jk}\partial^i - 2\delta^i_{(j}\partial_{k)}\right)\Phi \tag{3.44}$$

With these we can calculate our constraint equations from

$$\nabla_\mu T^\mu_\nu = \partial_\mu T^\mu_\nu + \Gamma^\mu_{\mu\alpha} T^\alpha_\nu - \Gamma^\alpha_{\mu\nu} T^\mu_\alpha = 0 \tag{3.45}$$

Our constrain equation for the $\nu = 0$ part is

$$\left(0^{\text{th}}\,\text{order}\right) \qquad \bar{\rho}' = -3\mathscr{H}\left(\bar{\rho} + \bar{P}\right) \tag{F3}$$
$$\left(1^{\text{st}}\,\text{order}\right) \qquad \delta\rho' = -3\mathscr{H}\left(\delta\rho + \delta P\right) + 3\Phi'\left(\bar{\rho} + \bar{P}\right) - \nabla\cdot\mathbf{q} \tag{3.46}$$

The three terms on the RHS. of (3.46) are
1. $3\mathscr{H}\left(\delta\rho + \delta P\right) \implies$ dilution due to expansion.
2. $3\Phi'\left(\bar{\rho} + \bar{P}\right) \implies$ perturbed expansion.
3. $\nabla\cdot\mathbf{q} \implies$ fluid flow.

If we use the definitions $\omega \equiv \frac{\bar{P}}{\bar{\rho}}$ and $c_s^2 \equiv \frac{\delta P}{\delta\rho}$ we can rewrite (3.46) as

Continuity Equation:

$$\delta' + (1+\omega)\left(\nabla\cdot\mathbf{v} - 3\Phi'\right) + 3\mathscr{H}\left(c_s^2 - \omega\right)\delta = 0 \tag{C}$$

Our constrain equation for the $\nu = i$ part is

$$\left(1^{\text{st}}\,\text{order}\right) \qquad \mathbf{v}' = -\mathscr{H}\mathbf{v} + 3\mathscr{H}\frac{\bar{P}'}{\bar{\rho}'}\mathbf{v} - \frac{\nabla\delta P}{\bar{\rho} + \bar{P}} - \nabla\Psi \tag{3.47}$$

The four terms on the RHS. of (3.47) are
1. $\mathscr{H}\mathbf{v} \implies$ redshift.
2. $3\mathscr{H}\frac{\bar{P}'}{\bar{\rho}'}\mathbf{v} \implies$ relativistic correction.
3. $\frac{\nabla\delta P}{\bar{\rho}+\bar{P}} \implies$ pressure gradients.
4. $\nabla\Psi \implies$ gravity.

Again using $\omega \equiv \frac{\bar{P}}{\bar{\rho}}$ and $c_s^2 \equiv \frac{\delta P}{\delta\rho}$, we can rewrite (3.47) as

Euler Equation:

$$\mathbf{v}' + \mathscr{H}\left(1 - 3\frac{\bar{P}'}{\bar{\rho}'}\right)\mathbf{v} = -\frac{c_s^2}{1+\omega}\nabla\delta - \nabla\Psi \tag{E}$$

These two equations are our constraint equations and hold for each fluid component separately. Now let us move onto the Einstein equations

We have

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\rho} \tag{3.48}$$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{3.49}$$

which if we substitute the expressions for the Christoffel symbols we eventually find that

$$G_{00} = 3\mathscr{H}^2 + 2\nabla^2\Phi - 6\mathscr{H}\Phi' \tag{3.50}$$

$$G_{0i} = 2\partial_i\left(\Phi' + \mathscr{H}\Psi\right) \tag{3.51}$$

$$\begin{aligned} G_{ij} = &-\left(2\mathscr{H}' + \mathscr{H}^2\right)\delta_{ij} + \partial_i\partial_j\left(\Phi - \Psi\right) \\ &+ \left[\nabla^2\left(\Psi - \Phi\right) + 2\Phi'' + 2\left(2\mathscr{H}' + \mathscr{H}^2\right)\left(\Phi + \Psi\right) + 2\mathscr{H}\Psi' + 4\mathscr{H}\Phi'\right]\delta_{ij} \end{aligned} \tag{3.52}$$

and combining with the expression for the energy momentum tensor we derive the following equations

First consider the trace free part of the ij equation

trace free ij part:

$$\left(1^{\text{st}}\,\text{order}\right) \quad \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)\left(\Phi - \Psi\right) = \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)\Pi$$

$$\implies \Phi - \Psi = \Pi \tag{G0}$$

so in the absence of anisotropic stress, $\Pi = 0$, the two Bardeen potential are equivalent $\Psi = \Phi$. We will assume this is the case from now on.

00 part:

$$\left(0^{\text{th}}\,\text{order}\right) \qquad 3\mathscr{H}^2 = 8\pi G a^2 \bar{\rho} \tag{F1}$$

$$\left(1^{\text{st}}\,\text{order}\right) \qquad \nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta + 3\mathscr{H}\left(\Phi' + \mathscr{H}\Phi\right) \tag{G1}$$

$0i$ part:

$$\left(1^{\text{st}}\,\text{order}\right) \quad \Phi' + \mathscr{H}\Phi = -4\pi G a^2 q \tag{G2}$$

ij trace part:

$$\left(0^{\text{th}}\,\text{order}\right) \qquad 2\mathscr{H}' + \mathscr{H}^2 = -8\pi G a^2 \bar{P} \tag{F2}$$

$$\left(1^{\text{st}}\,\text{order}\right) \quad \Phi'' + 3\mathscr{H}\Phi' + \left(2\mathscr{H}' + \mathscr{H}^2\right)\Phi = 4\pi G a^2 \delta P \tag{G3}$$

And combining (G1) and (G2) produces the Poisson equation:

Poisson Equation

$$\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\Delta \tag{P}$$

Which completes our set of equations. Before moving on to solving these equation we will first consider what type of perturbations we expect from inflation which will form our initial conditions

3.5 Adiabatic Perturbations

As we discussed at the end of our chapter on inflation the field ϕ moving down the potential can be viewed as a clock. The quantum fluctuations in ϕ can then be viewed a local time perturbation $\delta\tau$. This type of perturbation is called an adiabatic or curvature perturbation. It induces density and pressure perturbations in all fluids of the form

$$\delta\rho_I = \bar{\rho}_I(\tau+\delta\tau) - \bar{\rho}_I = \bar{\rho}_I'\delta\tau \tag{3.53}$$

and so we have

$$\delta\tau = \frac{\delta\rho_I}{\bar{\rho}_I'} = \frac{\delta\rho_J}{\bar{\rho}_J'} = \frac{\delta P_I}{\bar{P}_I'} \tag{3.54}$$

From this we can show

$$c_s^2 = \frac{\delta P}{\delta\rho} = \frac{\bar{P}'}{\bar{\rho}'} = \frac{d\bar{P}}{d\bar{\rho}} \implies P = P(\rho) \tag{3.55}$$

which is why they are called adiabatic perturbations. Also we note that for adiabatic perturbations with constant sound speed we have $c_s^2 = \omega$.

We can use the background continuity equation (F3) to relate

$$\bar{\rho}_I' \propto (1+\omega)\rho_I \implies \frac{\delta_I}{1+\omega_I} = \frac{\delta_J}{1+\omega_J} \tag{3.56}$$

so the density contrasts of all fluids are of similar size, with $\delta_r = \frac{4}{3}\delta_m$, and they all follow the same spatial profile, see Figure 3.1 for an example. This means that as

$$\delta\rho = \bar{\rho}\delta = \sum_I \bar{\rho}_I\delta_I \tag{3.57}$$

whatever dominates the background dominates the fluctuations.

Finally as we have shown that $\delta_I \propto \delta$ if we work in the constant density gauge $\delta = 0$ we set all density perturbations to zero for all fluids. The perturbation is now purely in the spatial curvature which is why they are called curvature perturbations.

3.6 Curvature Perturbation

Using gauge transformations we are free to move perturbations between time, density and 3-curvature. How do we measure the “real" perturbation?

As inflation predicts that we will have adiabatic modes sourced by a local time shift,

$$\delta\tau = \frac{\delta\rho}{\bar{\rho}'}\,, \tag{3.58}$$

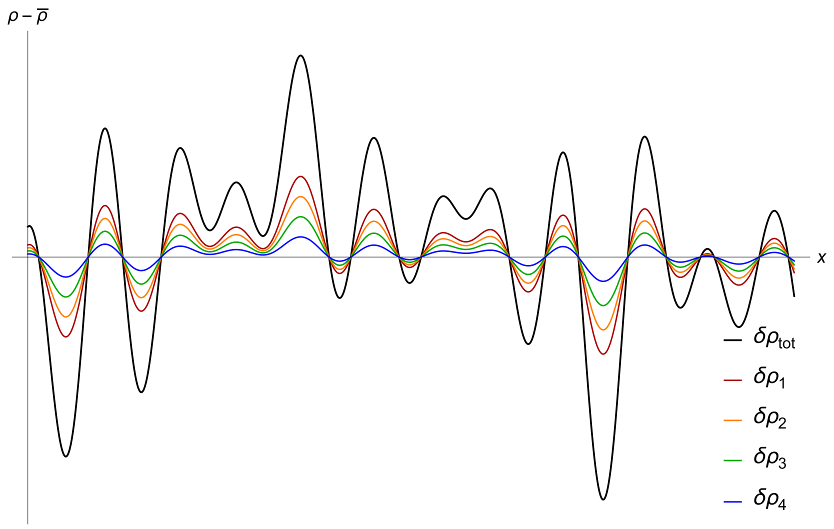


Figure 3.1: An example of Adiabatic perturbations, here we have four fluids and their total perturbation plotted together. The perturbations have different amplitudes but always the same profile.

if we take the uniform density gauge, $\delta\rho = B = 0$ then both $\delta\tau = \delta\rho = 0$, we have pushed all the perturbations into 3-curvature. Careful calculation reveals that the 3D Ricci scalar

$$a^2 R^{(3)} = 4\nabla^2 \left(-C + \frac{1}{3}\nabla^2 E\right) \tag{3.59}$$

so we will define the "constant density curvature perturbation" ζ to be

$$\zeta = \left[-C + \frac{1}{3}\nabla^2 E\right]_{\delta\rho=B=0} \tag{3.60}$$

This is useful but we would prefer it if it was gauge invariant. Under a gauge transformation

$$\begin{aligned}\tilde{\zeta} &= -\tilde{C} + \frac{1}{3}\nabla^2\tilde{E} = -C + \mathscr{H}T + \frac{1}{3}\nabla^2 L + \frac{1}{3}\nabla^2 (E-L) \\ &= \zeta + \mathscr{H}T\end{aligned} \tag{3.61}$$

so as $\tilde{\delta\rho} = \delta\rho - T\bar{\rho}'$ we can make ζ gauge invariant with the definition

Constant Density Curvature Perturbation

$$\zeta \equiv -C + \frac{1}{3}\nabla^2 E + \mathscr{H}\frac{\delta\rho}{\bar{\rho}'} \tag{3.62}$$

Now in spatially flat gauge $C = E = 0$ we see that

$$\zeta = \mathscr{H}\frac{\delta\rho}{\bar{\rho}'} = \mathscr{H}\delta\tau \tag{3.63}$$

and so it is directly related to the adiabatic perturbation

There is one other way to make ζ gauge invariant this time using B. This results in a similar quantity

$$\mathscr{R} \equiv -C + \frac{1}{3}\nabla^2 E - \mathscr{H}(B+v) \tag{3.64}$$

the **comoving curvature perturbation**. Both are popular and widely used (often with the opposite sign to that presented here) so what is the difference? Taking

$$\begin{aligned}\zeta - \mathscr{R} = \mathscr{H}\left(\frac{\delta\rho}{\bar{\rho}'} + B + v\right) &= \mathscr{H}\frac{\bar{\rho}}{\bar{\rho}'}\left(\frac{\delta\rho}{\bar{\rho}} + \frac{\bar{\rho}'}{\bar{\rho}}(v+B)\right) \\ &= \mathscr{H}\frac{\bar{\rho}}{\bar{\rho}'}\Delta \\ &= -\left(\frac{1}{3(1+\omega)}\right)\Delta \end{aligned} \tag{3.65}$$

so the difference is proportional to the comoving gauge density contrast Δ. Now (P) tells us that

$$\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\Delta = \frac{3}{2}\mathscr{H}^2\Delta \tag{3.66}$$

so in Fourier space we have $\Delta_k = \frac{2}{3}\left(\frac{k}{\mathscr{H}}\right)^2\Phi_k$ which on superhorizon scales ($k < \mathscr{H}$) is tiny. So on superhorizon scales we have

$$\zeta \approx \mathscr{R} \tag{3.67}$$

Aside: We are going to use the terms “superhorizon" and “subhorizon" alot in this and the next chapter (and in chapter 6) so it is worth pausing a moment to make sure we are clear what the terms mean.

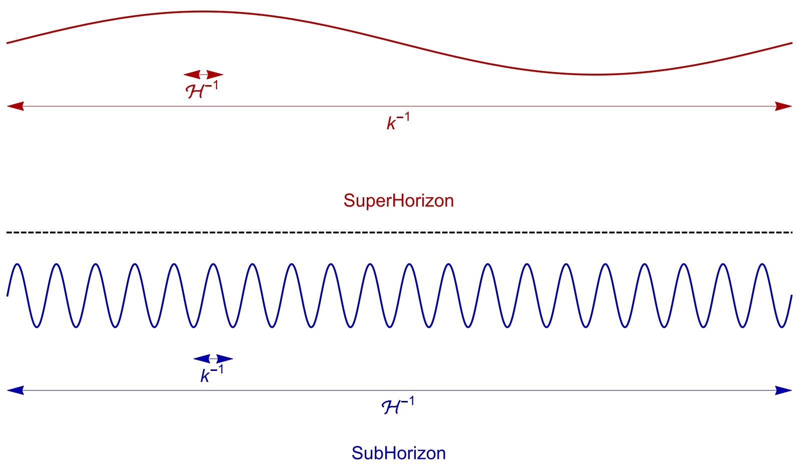


Figure 3.2: ‘A ‘Superhorizon" mode (top) which is much larger than the horizon Vs a “Subhorizon" mode (bottom) which is much smaller.

Superhorizon modes are ones whose wavelengths, k^{-1}, are much larger that the comoving Hubble radius, $\mathscr{H}^{-1}$, so $k \ll \mathscr{H}$. These modes have wavelengths that are much larger than the distance light can travel in one Hubble time. Thus the modes can be thought of as being out of causal contact with themselves and as a result can’t evolve dynamically. Instead, as gradients are effectively zero, they evolve by perturbing the local background so that locally for a quantity X the effective background $\bar{X}_{local}$ is

$$\bar{X}_{local} = \bar{X}_{global} + \delta X_{superhorizon} \tag{3.68}$$

The local causal patches then evolve like separate universes with perturbed backgrounds.

Subhorizon modes on the other hand are ones whose wavelengths, k^{-1}, are much smaller that the comoving Hubble radius, $\mathscr{H}^{-1}$, so $k \gg \mathscr{H}$. These modes are free to evolve dynamically in the usual way, see Figure 3.2

We have designed ζ to be the best gauge invariant measure of the adiabatic perturbation but it has one further property that makes it invaluable for structure formation. That is that ζ is conserved on superhorizon scales.

This property is sufficiently important that we will provide a full proof of it here. Consider ζ in Newtonian gauge (it can be proved in general but as it is gauge invariant if it is conserved in one it must be conserved in all)

$$\zeta = \Phi - \frac{1}{3}\frac{\delta\rho}{\bar{\rho}+\bar{P}} \tag{3.69}$$

Now take

$$\begin{aligned}
\left[3\left(\bar{\rho}+\bar{P}\right)\zeta\right]' &= \left[3\left(\bar{\rho}+\bar{P}\right)\Phi - \delta\rho\right]' \\
3\left(\bar{\rho}'+\bar{P}'\right)\zeta + 3\left(\bar{\rho}+\bar{P}\right)\zeta' &= 3\left(\bar{\rho}'+\bar{P}'\right)\Phi + 3\left(\bar{\rho}+\bar{P}\right)\Phi' - \delta\rho' \\
3\left(\bar{\rho}+\bar{P}\right)\zeta' &= 3\left(\bar{\rho}'+\bar{P}'\right)\left(\Phi-\zeta\right) + 3\left(\bar{\rho}+\bar{P}\right)\Phi' - \delta\rho'
\end{aligned} \tag{3.70}$$

Now we will use a form of (C),

$$\delta\rho' = -3\mathscr{H}\left(\delta\rho+\delta P\right) + 3\Phi'\left(\bar{\rho}+\bar{P}\right) - \nabla\cdot\mathbf{q} \tag{3.71}$$

to replace $\delta\rho'$ which removes the Φ' term and we can take care of the $\bar{\rho}'+\bar{P}'$ parts by considering

$$\begin{aligned}
3\left(\bar{\rho}'+\bar{P}'\right)\left(\Phi-\zeta\right) = 3\left(1+\frac{\bar{P}'}{\bar{\rho}'}\right)\bar{\rho}'\left(\Phi-\zeta\right) &= -3\mathscr{H}\left(1+\frac{\bar{P}'}{\bar{\rho}'}\right)3\left(\bar{\rho}+\bar{P}\right)\left(\Phi-\zeta\right) \\
&= -3\mathscr{H}\left(1+\frac{\bar{P}'}{\bar{\rho}'}\right)\delta\rho
\end{aligned} \tag{3.72}$$

where we have used (F3) to replace $\bar{\rho}'$ and the definition of ζ to get to the last line. Putting it together we have

$$\begin{aligned}
3\left(\bar{\rho}+\bar{P}\right)\zeta' &= -3\mathscr{H}\left(1+\frac{\bar{P}'}{\bar{\rho}'}\right)\delta\rho + 3\mathscr{H}\left(\delta\rho+\delta P\right) + \nabla\cdot\mathbf{q} \\
3\left(\bar{\rho}+\bar{P}\right)\zeta' &= 3\mathscr{H}\left(\delta P - \frac{\bar{P}'}{\bar{\rho}'}\delta\rho\right) + \nabla\cdot\mathbf{q} \\
\zeta' &= \mathscr{H}\frac{\delta P_{nad}}{\bar{\rho}+\bar{P}} + \frac{1}{3}\nabla\cdot\mathbf{v}
\end{aligned} \tag{3.73}$$

where on the last line we have defined the non-adiabatic pressure

$$\delta P_{nad} \equiv \delta P - \frac{\bar{P}'}{\bar{\rho}'}\delta\rho \tag{3.74}$$

which is clearly zero if $\frac{\delta P}{\delta\rho} = \frac{\bar{P}'}{\bar{\rho}'}$ which is true for adiabatic perturbations.

Now we can consider $\nabla\cdot\mathbf{v} = \nabla^2 v$. We can use (G2), $\Phi' + \mathscr{H}\Phi = \frac{3}{2}\mathscr{H}^2\left(1+\omega\right)v$, to show in Fourier space

$$\nabla^2 v_k \propto \left(\frac{k}{\mathscr{H}}\right)^2\left(\Phi' + \mathscr{H}\Phi\right) \approx 0 \quad \text{(superhorizon)} \tag{3.75}$$

So to summarise if our perturbations are adiabatic then on superhorizon scales we expect $\zeta =$ constant, ie: the adiabatic mode is conserved. This fact continues to hold at all orders for the suitably generalised form of ζ

Summary

In this chapter we derived the equations governing the evolution of perturbations. First we wrote down the most general perturbation of the FRW metric and performed a SVT decomposition breaking the perturbation into scalar, divergence-less vectors, and transverse-traceless tensor parts. At linear order the three decouple in the Einstein equations. Vectors have no obvious source and decay with expansion so we neglect them, tensors describe gravitational radiation which we will examine in the example sheets. This leaves us with just scalars which leaves the metric in the form

$$ds^2 = a^2(\tau)\left(-(1+2A)\,d\tau^2 + 2\partial_i B dx^i d\tau + \left[(1+2C)\,\delta_{ij} + 2\partial_{\langle i}\partial_{j\rangle}E\right]dx^i dx^j\right) \tag{3.76}$$

We do the same to the EM tensor producing

$$T_{\mu\nu} = a^2(\tau)\begin{pmatrix} \bar{\rho}+\delta\rho+2A\bar{\rho} & -\partial_i\left[q+\bar{\rho}B\right] \\ -\partial_i\left[q+\bar{\rho}B\right] & \bar{P}+\delta P+2C\bar{P}+\partial_{\langle i}\partial_{j\rangle}\left(\Pi+2\bar{P}E\right) \end{pmatrix} \tag{3.77}$$

where $q = (\bar{\rho}+\bar{P})\,v$ and $\partial_{\langle i}\partial_{j\rangle} = \partial_i\partial_j - \frac{1}{3}\nabla^2$ and we argued that Π is the anisotropic stress. We find that two of the perturbations are fake gauge modes related to coordinate transformations,

$$\mathbf{x}^\mu \to \widetilde{\mathbf{x}}^\mu = \mathbf{x}^\mu + \xi^\mu \tag{3.78}$$

and can either be removed via a gauge choice:

- Synchronous, $A = B = 0$
- Spatially Flat, $C = E = 0$
- Newtonian, $B = E = 0$
- Constant Density, $\delta\rho = B = 0$
- Comoving, $q = B = 0$

or by using gauge invariant quantities:

- $\Psi \equiv A + \mathscr{H}\left(B-E'\right) + \left(B-E'\right)'$
- $\Phi \equiv -C - \mathscr{H}\left(B-E'\right) + \frac{1}{3}\nabla^2 E$
- $\Delta \equiv \frac{\delta\rho}{\bar{\rho}} + \frac{\bar{\rho}'}{\bar{\rho}}\left(v+B\right)$

If we take the Newtonian gauge and substitute our expressions into the Einstein equations we calcualte both the 0^{th} order parts, which are the Freidmann equations, and the first order parts which give the following (with the definitions $\omega = P/\rho$ and $c_s^2 = \delta P/\delta\rho$):

$$\delta' + 3\mathscr{H}\left(c_s^2-\omega\right)\delta = (1+\omega)\left(3\Phi' - \nabla\cdot\mathbf{v}\right) \tag{C}$$

$$\mathbf{v}' + \mathscr{H}\left(1-3\frac{\bar{P}'}{\bar{\rho}'}\right)\mathbf{v} = -\frac{c_s^2}{1+\omega}\nabla\delta - \nabla\Psi \tag{E}$$

$$\nabla^2\Phi - 3\mathscr{H}\left(\Phi' + \mathscr{H}\Phi\right) = 4\pi G a^2\bar{\rho}\delta \tag{G1}$$

$$\Phi' + \mathscr{H}\Phi = -4\pi G a^2 q \tag{G2}$$

$$\Phi'' + 3\mathscr{H}\Phi' + \left(2\mathscr{H}' + \mathscr{H}^2\right)\Phi' = 4\pi G a^2\delta P \tag{G3}$$

$$\implies \nabla^2\Phi = 4\pi G a^2\bar{\rho}\Delta \tag{P}$$

We defined adiabatic perturbations which are those sourced by a local perturbation in time

$$\delta\tau = \frac{\delta\rho}{\bar{\rho}'} = \frac{\delta P}{\bar{P}'} \tag{3.79}$$

This type of perturbation affects all constituent fluids equally so they have similar density contrasts $\delta_I \approx \delta_J$ and the perturbations all have the same spatial profile.

Finally we defined the best measure of the adiabatic perturbation, the constant density curvature perturbation ζ, and the closely related quantity, the comoving curvature perturbation $\mathscr{R}$.

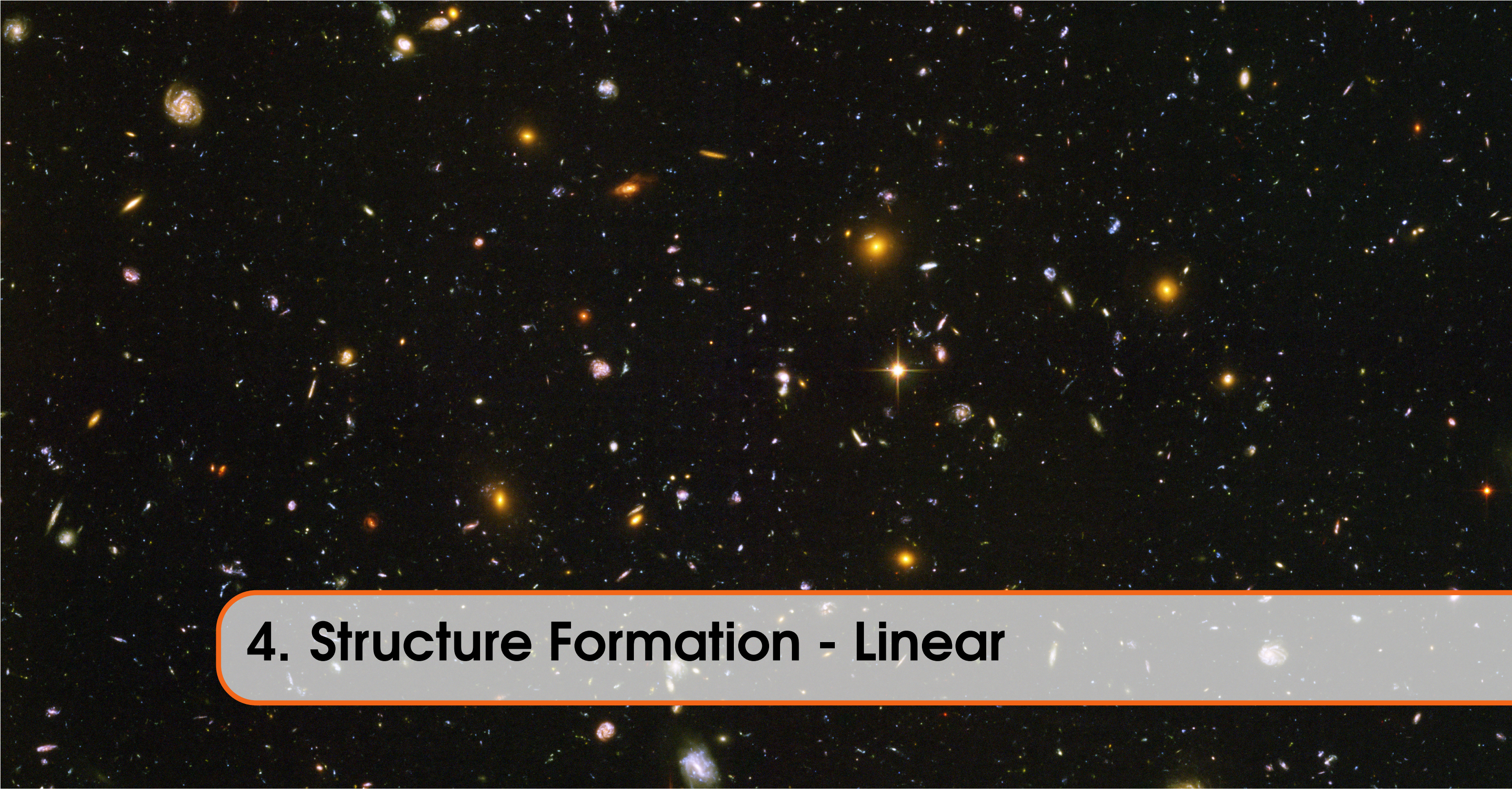
$$\zeta \equiv -C + \frac{1}{3}\nabla^2 E + \mathscr{H}\frac{\delta\rho}{\bar{\rho}'} \tag{3.80}$$

$$\mathscr{R} \equiv -C + \frac{1}{3}\nabla^2 E - \mathscr{H}(B+v) \tag{3.81}$$

We showed that their difference was

$$\zeta - \mathscr{R} = -\left(\frac{1}{3(1+\omega)}\right)\Delta \tag{3.82}$$

which is ≈ 0 on superhorizon scales. We also proved that ζ (and by association $\mathscr{R}$) are conserved on superhorizon scales when the initial perturbation is adiabatic.



4. Structure Formation - Linear

Here we will solve the linear perturbation equations for both the radiation and matter perturbations for all eras. Let us first summarise the equations for adiabatic perturbations.

Background Equations:

$$3\mathscr{H}^2 = 8\pi G a^2 \bar{\rho} \tag{F1}$$

$$2\mathscr{H}' + \mathscr{H}^2 = -8\pi G a^2 \bar{P} \tag{F2}$$

$$\implies \bar{\rho}' = -3\mathscr{H}\left(\bar{\rho} + \bar{P}\right) \tag{F3}$$

Evolution Equations:

$$\nabla^2\Phi - 3\mathscr{H}\left(\Phi' + \mathscr{H}\Phi\right) = 4\pi G a^2 \bar{\rho}\delta = \tfrac{3}{2}\mathscr{H}^2\delta \tag{G1}$$

$$\Phi' + \mathscr{H}\Phi = -4\pi G a^2 q = -\tfrac{3}{2}\mathscr{H}^2\left(1+\omega\right)v \tag{G2}$$

$$\Phi'' + 3\mathscr{H}\Phi' + \left(2\mathscr{H}' + \mathscr{H}^2\right)\Phi = 4\pi G a^2 \delta P \tag{G3}$$

$$\implies \nabla^2\Phi = 4\pi G a^2 \bar{\rho}\Delta = \tfrac{3}{2}\mathscr{H}^2\Delta \tag{}$$

Constraint Equations:

$$\delta' = \left(1+\omega\right)\left(3\Phi' - \nabla\cdot\mathbf{v}\right) \tag{C}$$

$$\mathbf{v}' + \mathscr{H}\left(1-3\omega\right)\mathbf{v} = -\frac{\omega}{1+\omega}\nabla\delta - \nabla\Phi \tag{E}$$

4.1 Initial Conditions

If we take (G3)$-\omega$(G1) then we obtain a closed equation for the potential

$$\Phi'' + 3(1+\omega)\mathscr{H}\Phi' - \omega\nabla^2\Phi = 0 \tag{4.1}$$

On superhorizon scales we can neglect gradients so the growing mode is

$$\Phi = \text{const.} \tag{4.2}$$

for all eras. Now we can take (G1) neglecting all derivatives to find

$$-3\mathscr{H}^2\Phi = \tfrac{3}{2}\mathscr{H}^2\delta \implies \delta = -2\Phi \tag{4.3}$$

so δ is also constant on superhorizon scales (As the perturbations are adiabatic this is true for all fluid components).

Aside: But using (P)

$$\nabla^2\Phi = \frac{3}{2}\mathscr{H}^2\Delta \implies \Delta_k = -\frac{2}{3}\left(\frac{k}{\mathscr{H}}\right)^2\Phi_k \tag{4.4}$$

where the last step involves switching to Fourier space. So on superhorizon scales $\Delta \ll \delta$, however now Δ isn't constant but is growing as $\mathscr{H}^{-2} \propto \tau^2$. This seems odd as both Δ and δ are supposed to be measures of the same thing, the density contrast, but in different gauges. This is a physical quantity so we can go out and measure it. Which one is correct? The answer is that both are as you can't make observations on superhorizon scales. On subhorizon scales we have

$$\Delta - \delta = -3\mathscr{H}(1+\omega)v \propto \left(\frac{\mathscr{H}}{k}\right)\mathbf{v} \approx 0 \tag{4.5}$$

so on subhorizon scales, where experiments are possible, the two are the same and there is no gauge ambiguity. This is a general result, for physical quantities the Gauge problem always vanishes on subhorizon scales.

Now from the definition of ζ we have

$$\zeta = \Phi - \frac{1}{3(1+\omega)}\delta = \frac{5+3\omega}{3+3\omega}\Phi \tag{4.6}$$

So as we know ζ is conserved on superhorizon scales we have initial conditions of

$$\zeta = \frac{3}{2}\Phi_{RD} = \frac{5}{3}\Phi_{MD} \tag{4.7}$$

where the subscript RD stands for radiation domination and the subscript MD stands for matter domination. One interesting consequence of this is that

$$\Phi_{MD} = \frac{9}{10}\Phi_{RD} \tag{4.8}$$

So while Φ is conserved on superhorizon scales when the equation of state is constant this isn't true when it varies between eras. What happens here is that the decaying modes is excited during transition and some of the potential leaks away. Now we have our initial conditions from ζ in terms of the potential Φ so we can solve the dynamics.

4.2 Potential ϕ

We already know the solution is $\Phi = \text{const.}$ on superhorizon scales with a rescaling by $9/10$ when we transition from radiation domination to matter domination from the previous section. But what about on subhorizon scales in the various eras?

4.2.1 Radiation Domination

During radiation domination (4.1) has the form in Fourier space

$$\Phi_k'' + \frac{4}{\tau}\Phi_k' + \frac{k^2}{3}\Phi_k = 0 \tag{4.9}$$

Which has solution

$$\Phi_k(\tau) = 2\zeta_k(0)\left(\frac{\sin(x) - x\cos(x)}{x^3}\right), \quad x = \frac{k\tau}{\sqrt{3}} \tag{4.10}$$

On superhorizon scales when $x \ll 1$ we have

$$\Phi_k^{super}(\tau) = \frac{2}{3}\zeta_k(0) \tag{4.11}$$

Which agrees with the solution we found previously. On subhorizon scales where $x \gg 1$ then we have

$$\Phi_k^{sub}(\tau) = 6\zeta_k(0)\frac{\cos(x)}{x^2} \propto \frac{\text{“Oscillations"}}{a^2} \tag{4.12}$$

4.2.2 Matter Domination

During matter domination (4.1) has the form

$$\Phi_k'' + \frac{6}{\tau}\Phi_k' = 0 \tag{4.13}$$

so the solution is

$$\Phi_k(\tau) = \frac{3}{5}\zeta_k(0)\,. \tag{4.14}$$

We see that during matter domination Φ is constant on all scales. We summarise the evolution of Φ in Figure 4.1

4.3 Radiation Perturbation

On superhorizon scales we know δ_r is constant and Δ_r grows as τ^2 but what happens on subhorizon scales?

4.3.1 Radiation Domination

First we will calculate the evolution of the radiation perturbation during radiation domination. As the radiation is dominant we can use (P)

$$\nabla^2\Phi = \frac{3}{2}\mathscr{H}^2\Delta_r \tag{4.15}$$

In Fourier space this implies

$$\Delta_r(k) = -\frac{2}{3}\left(\frac{k}{\mathscr{H}}\right)^2\Phi_k \propto (k\tau)^2\Phi_k \propto a^2\Phi_k \tag{4.16}$$

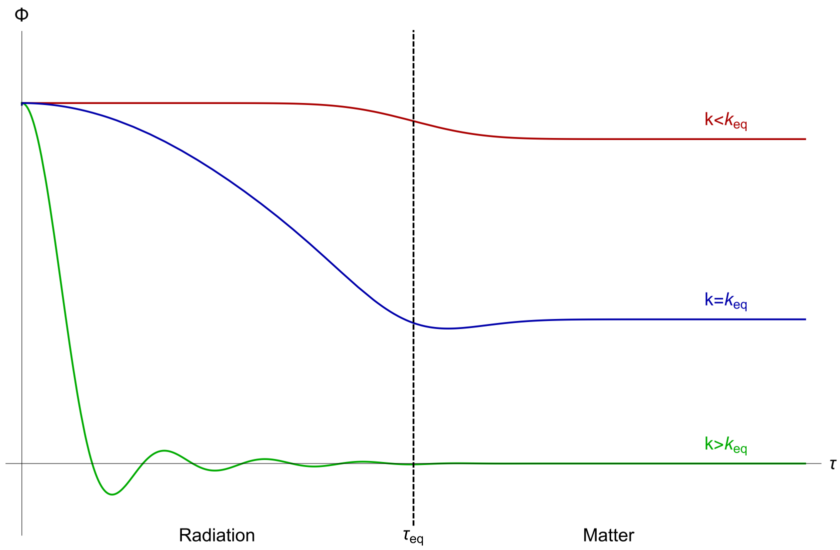


Figure 4.1: Evolution of the potential Φ during matter and radiation eras

So in the radiation era we can just take our potential solutions and multiply them by a^2. So

$$\text{Superhorizon} \implies \Delta_r \propto a^2, \quad \delta_r = \text{const.} \tag{4.17}$$

$$\text{Subhorizon} \implies \Delta_r \approx \delta_r \propto \cos\left(\frac{k\tau}{\sqrt{3}}\right) \tag{4.18}$$

so on subhorizon scales the radiation just oscillates. This is just the effect of the two opposing forces acting on the photon fluid, gravity and radiation pressure. Perturbations collapse under gravity which increases the radiation pressure which eventually overcomes gravity and causes expansion. Pressure then drops and gravity takes over again forcing collapse. This is the basic mechanism behind the oscillations we see both in the cosmic microwave background and in baryon acoustic oscillations.

4.3.2 Matter Domination

Now the radiation perturbation is subdominant so we have to use the constraint equations, (C)+(E), to determine its evolution. The potential is constant during matter domination so the constrain equations for radiation are

$$\delta_r' = -\frac{4}{3}\nabla \cdot \mathbf{v}_r \tag{4.19}$$

$$\mathbf{v}_r' = -\frac{1}{4}\nabla \delta_r - \nabla \Phi_{MD} \tag{4.20}$$

where the subscript MD remind us to use the matter domination solution. These 2 equations combine to form

$$\delta_r'' - \frac{1}{3}\nabla^2 \delta_r = \frac{4}{3}\nabla^2 \Phi_{MD} \tag{4.21}$$

This is the equation of an oscillator with constant force. The solution is

$$\delta_r(k) = -4\Phi_{MD} + \text{“Oscillations"} \tag{4.22}$$

So during the matter era the radiation perturbation oscillates around the potential sourced by matter perturbations. We summarise the evolution of Δ_r in Figure 4.2. We should note that this analysis does not take into account the very significant effect of coupling to baryons. This is addressed properly in the Advanced Cosmology course next term when you derive the equations governing the cosmic microwave background.

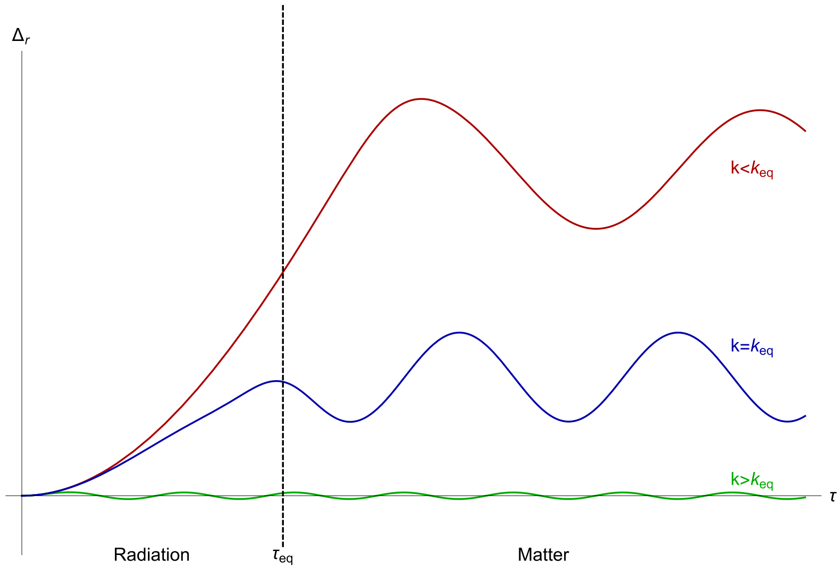


Figure 4.2: Evolution of the comoving radiation density contrast Δ_r during matter and radiation eras

4.4 Matter Perturbation

Now let's consider the matter perturbation. Again on superhorizon scales we know δ_m is constant and Δ_m grows as τ^2. What happens on subhorizon scales?

4.4.1 Matter Domination

As matter is the dominant component we can use (P) to see that

$$\Delta_m(k) = -\frac{2}{3}\left(\frac{k}{\mathscr{H}}\right)^2 \Phi_k = -\frac{1}{6}(k\tau)^2 \Phi_k \propto a(\tau) \tag{4.23}$$

where this is now true on all scales.

4.4.2 Radiation + Matter Eras

As matter is not always the dominant component we need to deduce its evolution from the constraint equations sourced by the potential. For matter the constraint equations (C)+(E) are

$$\delta_m' = -\nabla \cdot \mathbf{v}_m + 3\Phi' \tag{4.24}$$

$$\mathbf{v}_m' = -\mathscr{H}\mathbf{v}_m - \nabla\Phi \tag{4.25}$$

where Φ is sourced by both the matter and radiation so $\Phi = \Phi_r + \Phi_m$. We can combine these into one equation

$$\delta_m'' + \mathscr{H}\delta_m' = \nabla^2\Phi + 3\left(\Phi'' + \mathscr{H}\Phi'\right) \tag{4.26}$$

Now on subhorizon scales we know that

$$\Phi_r \propto \cos(k\tau) \approx \cos(\frac{k}{\mathscr{H}}) \tag{4.27}$$

$$\Phi_m \approx \text{const.} \tag{4.28}$$

So the radiation part is rapidly oscillating. The result is that both δ_r and δ_m have two modes

- A “fast" mode sourced by the radiation
- A “slow" mode sourced by the matter

The radiation perturbation is comfortable with the rapid oscillations and the “fast" mode generaly dominates. However, as matter has both mass and zero pressure oscillatory solutions are somewhat unnatural and “fast" solutions are suppressed by a factor $(\mathscr{H}/k)^2$ relative to “slow" modes. In effect the matter only sees the time average of the gravitational potential as it can't react very fast to change. The result is that the matter perturbation is soured by the matter potential even deep in the radiation era. This allows us to make the following simplifications when considering the matter perturbation

$$\Phi \approx \Phi_m \tag{4.29}$$

$$\Phi' \approx \Phi'' \approx 0 \tag{4.30}$$

$$\nabla^2 \Phi_m = \frac{3}{2}\mathscr{H}^2 \Delta_m \approx \frac{3}{2}\mathscr{H}^2 \delta_m \tag{4.31}$$

Using these with (4.26) we obtain

$$\delta_m'' + \mathscr{H}\delta_m' - \frac{3}{2}\mathscr{H}^2 \delta_m = 0 \tag{4.32}$$

and we just need an expression for $\mathscr{H}$ in both eras.

Exercise: Show that for a universe with only radiation and matter (F1) can be written as

$$\mathscr{H}^2 = \frac{H_0^2 \Omega_m^2}{\Omega_r}\left(\frac{1}{y} + \frac{1}{y^2}\right), \quad y \equiv \frac{a}{a_{eq}} \tag{4.33}$$

where a_{eq} is the scale factor at the time of matter radiation equality.

With this we can substitute for $\mathscr{H}$ and we get the

Mészáros Equation:

$$\frac{d^2\delta_m}{dy^2} + \frac{2+3y}{2y(1+y)}\frac{d\delta_m}{dy} - \frac{3}{2y(1+y)}\delta_m = 0 \tag{M}$$

Which has solutions

$$\delta_m \propto \begin{cases} 2+3y \\ (2+3y)\ln\left(\frac{\sqrt{1+y}+1}{\sqrt{1+y}-1}\right) - 6\sqrt{1+y} \end{cases} \tag{4.34}$$

During radiation domination $y \ll 1$ so solutions are

$$\delta_m \propto \ln(y) \propto \ln(a) \propto \ln(\tau) \tag{4.35}$$

And during matter domination $y \gg 1$ so solutions are

$$\delta_m \propto y \propto a \propto \tau^2 \tag{4.36}$$

as before. The result is that the matter perturbation grows as τ^2 everywhere except on sub horizon scales during radiation domination where it only grows as $\ln(\tau)$. This is the Mészáros effect which is important calculating the matter power spectrum as we shall see later. The physical interpretation is that matter perturbations grow everywhere under gravitational collapse except on subhorizon scales during radiation domination where the rapid oscillation of the radiation fluid stops gravitational wells from forming and the matter perturbation stalls.

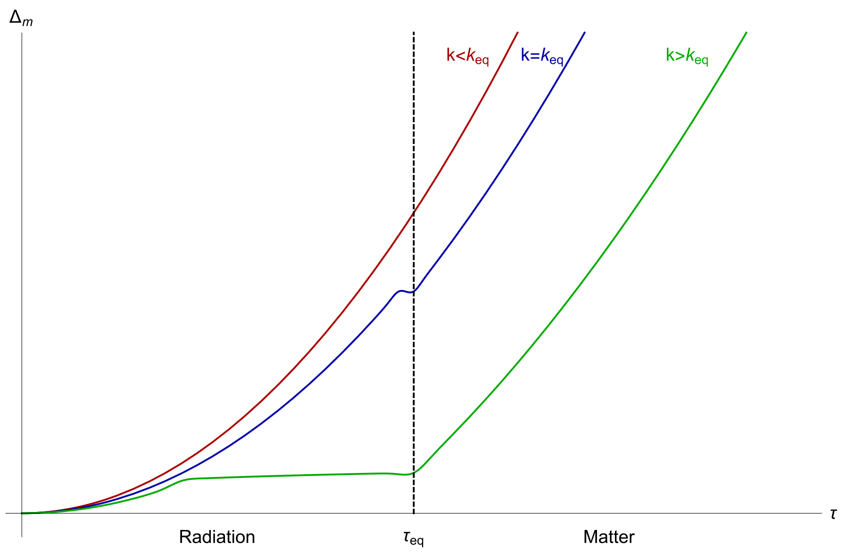


Figure 4.3: Evolution of the comoving matter density contrast Δ_m during matter and radiation eras

4.5 Baryon Perturbation

Before moving onto power spectra we will quickly look at baryons. Up to now we have treated baryons and cold dark matter the same but there are two crucial differences. The first is that baryons have pressure the second is that before electrons bind to from neutral hydrogen at recombination, the baryons interact with photons via Thomson scattering.

4.5.1 Radiation Domination

During radiation domination all electrons are free and so the baryon and radiation fluids are tightly coupled via Thompson scattering. This means that unlike the matter which ignores the oscillations in radiation the baryons follow the rapid oscillations of the radiation. So on subhorizon scales,

$$\Delta_b \approx \Delta_\gamma \approx \Delta_r \propto \cos\left(\frac{k\tau}{\sqrt{3}}\right). \tag{4.37}$$

4.5.2 Matter Domination

Recombination happens shortly after matter-radiation equality so during matter domination the baryons decouple from the radiation fluid and behave as matter but with non-zero pressure. We can examine the behaviour of the baryons from the constraint equations (C)+(E). For both baryons and cold dark matter they can be written as

$$\delta_b'' + \mathscr{H}\delta_b' - c_s^2\nabla^2\delta_b = \nabla^2\Phi \tag{4.38}$$

$$\delta_c'' + \mathscr{H}\delta_c' = \nabla^2\Phi \tag{4.39}$$

If we define the fractional difference $\delta_{bc} = \delta_b - \delta_c$ then we can take the difference of these two equations to find (in Fourier space)

$$\delta_{bc}''(k) + \mathscr{H}\delta_{bc}'(k) = -c_s^2 k^2 \delta_b(k) \tag{4.40}$$

For $k \ll c_s$ the solution is $\delta_{bc} \approx$ const. so during matter domination we have

$$\frac{\delta_{bc}}{\delta_m} \propto \frac{1}{a} \tag{4.41}$$

and the proportional difference decays. From this we can deduce that on scales larger than the sound horizon the baryons follow the matter. What about on scales smaller than the sound horizon? If we take the constraint equation for the baryons and use $\nabla^2\Phi = \frac{3}{2}\mathscr{H}^2\delta_m \approx \frac{3}{2}\mathscr{H}^2\delta_b$ then we have

$$\delta_b'' + \mathscr{H}\delta_b' - c_s^2\nabla^2\delta_b = \frac{3}{2}\mathscr{H}^2\delta_b \tag{4.42}$$

$$\delta_b'' + \mathscr{H}\delta_b' + \left(c_s^2 k^2 - \frac{3}{2}\mathscr{H}^2\right)\delta_b = 0 \tag{4.43}$$

From this we can see that if

$$k > \sqrt{\frac{3}{2}}\frac{\mathscr{H}}{c_s} \tag{4.44}$$

then neglecting the friction term, $\mathscr{H}\delta_b'$, we have a wave equation and the baryons oscillate. The physical length scale, $\lambda_J = a(2\pi/k)$, when we cross over into this oscillatory behaviour is known as the Jeans length and using (F1) to replace $\mathscr{H}$ we have

Jeans Length:

$$\lambda_J = c_s\sqrt{\frac{\pi}{G\bar{\rho}}} \tag{J}$$

This is the smallest scale at which we can have collapse. Above this scale the baryons follow the dark matter and structures grow. Below it the baryon perturbations are pressure supported and oscillate.

4.5.3 Baryon Acoustic Oscillations

What is the effect of the baryons switching from following the radiation to following the matter? On scales above the Jeans length at late times we have all matter perturbations being approximately the same. As matter perturbations grow as a during matter domination we can relate them to an earlier time by the ratio of the scale factors:

$$\Delta_c(\tau_0) \approx \Delta_b(\tau_0) \approx \Delta_m(\tau_0) \approx \Delta_m(\tau^*)\left(\frac{a(\tau_0)}{a(\tau^*)}\right). \tag{4.45}$$

Now if we take the time τ^* to be the time of recombination when the baryons were just decoupled from the photons then we have

$$\Delta_m = \Delta_m^* \frac{a_0}{a^*} = (f_b \Delta_b^* + f_c \Delta_c^*) \frac{a_0}{a^*}, \quad f_x \equiv \frac{\bar{\rho}_x}{\bar{\rho}} \tag{4.46}$$

Now at recombination, which is early in the matter era, we have

$$\Delta_b^* \approx \Delta_\gamma^* \propto -4\Phi_k + A\cos\left(\frac{k\tau^*}{\sqrt{3}}\right) \tag{4.47}$$

So we find that the matter perturbation at late times still has a small imprint of the oscillations present in the baryons at the time of recombination,

$$\Delta_m \approx \frac{1}{f_c}\Delta_c\left(1 + \varepsilon\cos\left(\frac{k\tau^*}{\sqrt{3}}\right)\right) \tag{4.48}$$

where ε is some small amplitude. These small oscillations can be seen in the matter distribution today and are called the Baryon Acoustic Oscillations (BAO). They are important as measuring their frequency gives a measure of τ^* the time of recombination. Here is a table summarising the

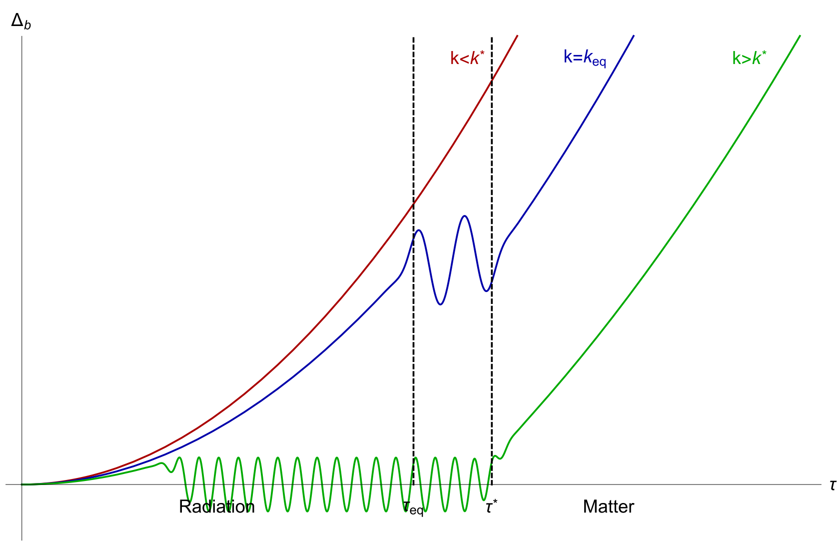


Figure 4.4: Evolution of the comoving baryon density contrast Δ_b during matter and radiation eras

solutions we have found

Summary of the evolution of perturbations

	Radiation Domination	Matter Domination
Potential Φ		
– Superhorizon	const.	const.
– Subhorizon	$\propto a^{-2}\cos\left(\frac{k\tau}{\sqrt{3}}\right)$	const.
Radiation Δ_r (δ_r)		
– Superhorizon	$\propto a^2 \propto \tau^2$ (const.)	$\propto a \propto \tau^2$ (const.)
– Subhorizon	$\propto \cos\left(\frac{k\tau}{\sqrt{3}}\right)$	$\propto$ const. + oscillations
Matter Δ_m (δ_m)		
– Superhorizon	$\propto a^2 \propto \tau^2$ (const.)	$\propto a \propto \tau^2$ (const.)
– Subhorizon	$\propto \ln(a) \propto \ln(\tau)$	$\propto a \propto \tau^2$
Baryons Δ_b (δ_b)		
– Superhorizon	$\propto a^2 \propto \tau^2$ (const.)	$\propto a \propto \tau^2$ (const.)
– Subhorizon	$\propto \cos\left(\frac{k\tau}{\sqrt{3}}\right)$	$\propto a \propto \tau^2$

4.6 Power Spectra

Now we know the evolution of all fluid perturbations in all eras. How do we measure them? The perturbations are Gaussian (as the primordial perturbations are Gaussian and we are in the linear regime) so are completely described by their mean (which is zero by definition) and their power spectrum, P defined by,

$$\left\langle \Delta_x(\mathbf{k})\Delta_x(\mathbf{k}')\right\rangle = (2\pi)^3\,\delta\left(\mathbf{k}-\mathbf{k}'\right)P_x(k) \tag{4.49}$$

The power spectrum is just the variance on scale k.

4.6.1 Initial Conditions

We know that on superhorizon scales that $\delta \propto \Phi \propto \zeta$ and we expect a scale invariant spectrum of perturbations in ζ from inflation which means

$$P_\delta \propto P_\Phi \propto P_\zeta \propto k^{n_s-4} \tag{4.50}$$

On superhorizon scales we also have via (P)

$$\Delta \propto \left(\frac{k}{\mathscr{H}}\right)^2 \Phi_k \implies P_\Delta(\tau_i) \propto k^4 P_\Phi(\tau_i) \propto k^{n_s} \tag{4.51}$$

Where we have to do it at a fixed time as Δ is growing as τ^2 on superhorizon scales. Thus our initial power spectra at some initial time τ_i when all modes are inside the horizon are

$$P_\delta(\tau_i,k) \propto k^{n_s-4} \tag{4.52}$$
$$P_\Delta(\tau_i,k) \propto k^{n_s} \tag{4.53}$$

4.6.2 Matter Power Spectrum

If we take the results for the growth of matter perturbations an plot them in the k-τ plane it becomes clear that for all regions we are interested in the matter perturbation, Δ_m, grows as τ^2 except for during radiation domination inside the horizon where the Mészáros effect occurs and growth stalls, see Figure 4.5.

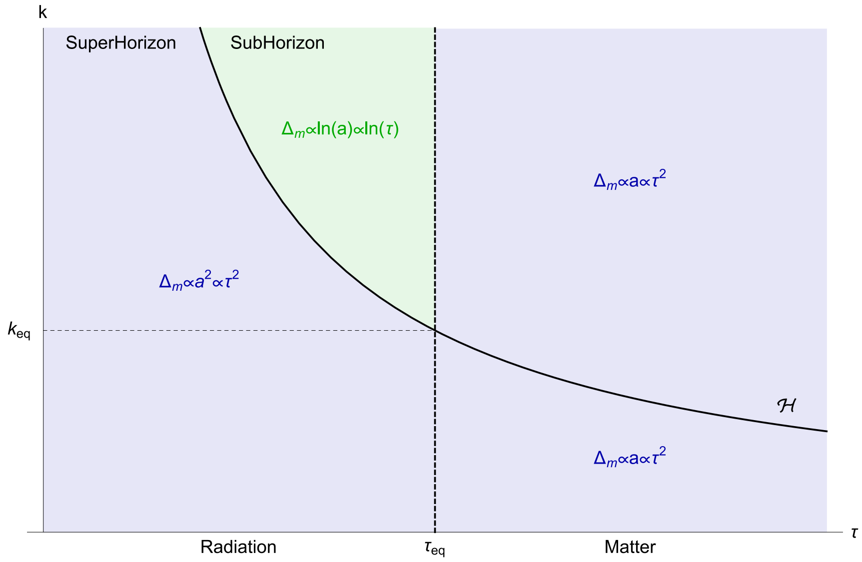


Figure 4.5: Evolution of the matter perturbation

From this we see that for scales $k < k_{eq}$ the primordial perturbation grows by a factor $(\tau_0/\tau_i)^2$ so

$$P_{\Delta_m}(\tau_0, k < k_{eq}) = \left(\frac{\tau_0}{\tau_i}\right)^4 P_{\Delta_m}(\tau_i, k < k_{eq}) \propto k^{n_s} \tag{4.54}$$

However for scales which became subhorizon during radiation we must cross the Mészáros region where growth stalls. Now if we define τ_X as the time of horizon crossing, we find

$$\begin{aligned} P_{\Delta_m}(\tau_0, k > k_{eq}) &= \left(\frac{\tau_0}{\tau_{eq}}\right)^4 \left(1 + \ln\left(\frac{\tau_{eq}}{\tau_X}\right)\right)^2 \left(\frac{\tau_X}{\tau_i}\right)^4 P_{\Delta_m}(\tau_i, k > k_{eq}) & (4.55) \\ &= \left(\frac{\tau_0}{\tau_{eq}}\right)^4 \left(1 + \ln\left(\frac{k}{k_{eq}}\right)\right)^2 \left(\frac{1}{k\tau_i}\right)^4 P_{\Delta_m}(\tau_i, k > k_{eq}) \propto \ln^2(k)\, k^{n_s - 4} & (4.56) \end{aligned}$$

where we have used that at horizon crossing $k \approx \mathscr{H} \propto \tau^{-1}$ so $\tau_X \approx 1/k$. Finally these modes have a BAO component so we should add a $\left(1 + 2\varepsilon\cos(k\tau^*/\sqrt{3})\right)^2$ factor. The power matter power spectrum has the following form

$$P_{\Delta_m}(\tau_0, k) = A_{\Delta_m} \begin{cases} \left(\frac{k}{k_{eq}}\right)^{n_s} & k < k_{eq} \\ \left(1 + 2\varepsilon\cos\left(k\tau^*/\sqrt{3}\right)\right)^2 \times \left(1 + \ln\left(\frac{k}{k_{eq}}\right)\right)^2 \times \left(\frac{k}{k_{eq}}\right)^{n_s - 4} & k > k_{eq} \end{cases} \tag{4.57}$$

4.6.3 Radiation Power Spectrum

We can apply the same arguments of the previous section to get a basic picture of the power spectrum of the CMB. We have neglected completely the interactions of the photons with baryons so the picture is extremely crude but let us continue. We have three regions to consider. On superhorizon scales Δ_r grow as τ^2, on subhorizon scales during radiation domination the perturbation oscillates and on subhorizon scales during matter domination the perturbations are roughly constant (the

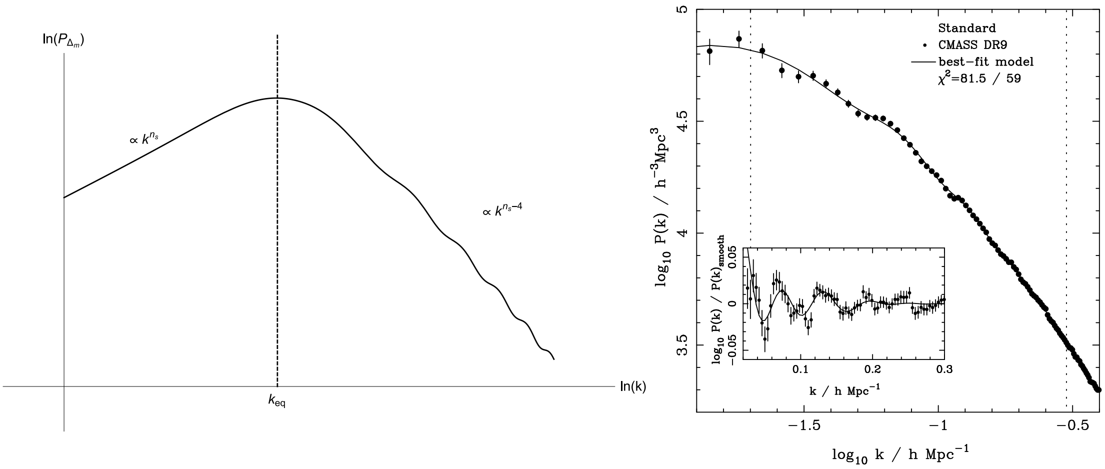


Figure 4.6: Matter Power Spectrum. Left is our prediction. Right is observations from SDSS III with BAO shown in inset. Scales above k_{eq}^{-1} are too large to be easily observed with Galaxies and constraints instead come from the CMB (not shown)

.

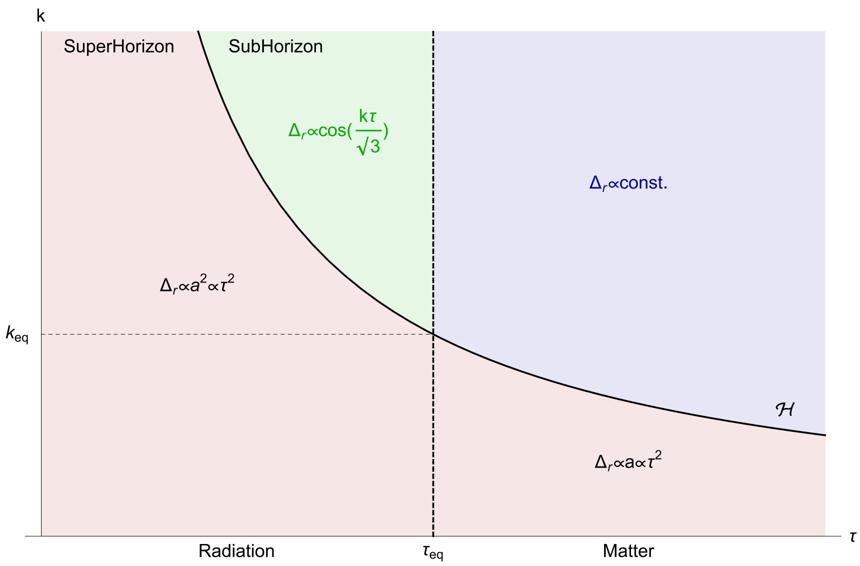


Figure 4.7: Evolution of the radiation perturbation

oscillations are much smaller than Φ). These can be summarised in Figure 4.7 From this we see that for scales $k < k_{eq}$ the primordial perturbation grows by a factor $(\tau_X/\tau_i)^2$ so

$$P_{\Delta_r}(\tau_0, k < k_{eq}) = \left(\frac{\tau_X}{\tau_i}\right)^4 P_{\Delta_r}(\tau_i, k < k_{eq}) \propto k^{n_s - 4} \tag{4.58}$$

However for scales which became subhorizon during radiation we must cross the oscillating region so we have

$$P_{\Delta_r}(\tau_0, k > k_{eq}) = \left(\frac{\tau_X}{\tau_i}\right)^4 \cos^2\left(\frac{k\tau_{eq}}{\sqrt{3}}\right) P_{\Delta_r}(\tau_i, k > k_{eq}) \propto k^{n_s - 4} \cos^2\left(\frac{k\tau_{eq}}{\sqrt{3}}\right) \tag{4.59}$$

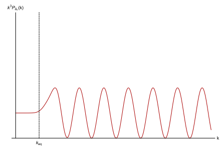
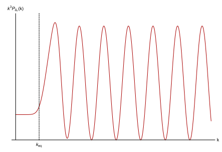
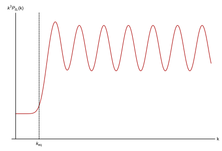
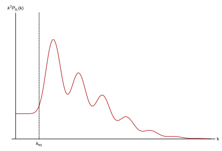
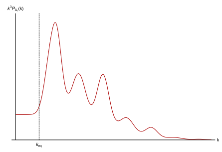
1. The naive prediction for the radiation power spectrum neglecting baryons	
2. When we couple baryons to the photons they increase the gravitational force on the radiation which increases the amplitude of the oscillations	
3. We also have a velocity component to the power spectrum from the Doppler effect (regions which are moving towards us seem hotter, those moving away seem colder) which adds a subdominant $\sin^2(k)$ term	
4. Next we need to account for the fact that the radiation and baryons are not perfectly coupled so the photons tend to diffuse out of overdense regions damping the oscillations. This leads to an exponential suppression in power for large k	
5. Finally we need to note that there is less diffusion during compression (maximum density) and more during expansion expansion (minimum density) this increases the height of odd peaks versus even peaks (the first thing an overdensity does on entering the horizon is collapse so the first peak is compressive)	

Table 4.1: Radiation Power Spectrum

The approximate radiation power spectrum has the following form

$$P_{\Delta_r}(\tau_0,k) = A_{\Delta_m}\begin{cases}\left(\frac{k}{k_{eq}}\right)^{n_s-4} & k<k_{eq}\\ \cos^2\left(\frac{k\tau_{eq}}{\sqrt{3}}\right)\times\left(\frac{k}{k_{eq}}\right)^{n_s-4} & k>k_{eq}\end{cases} \tag{4.60}$$

which is scale invariant with oscillations for $k > k_{eq}$. This is very crude however we can make an improved qualitative estimate with some simple modifications, see Table 4.1.

The result of this is a power spectrum which is close to the one observed, see Figure 4.8

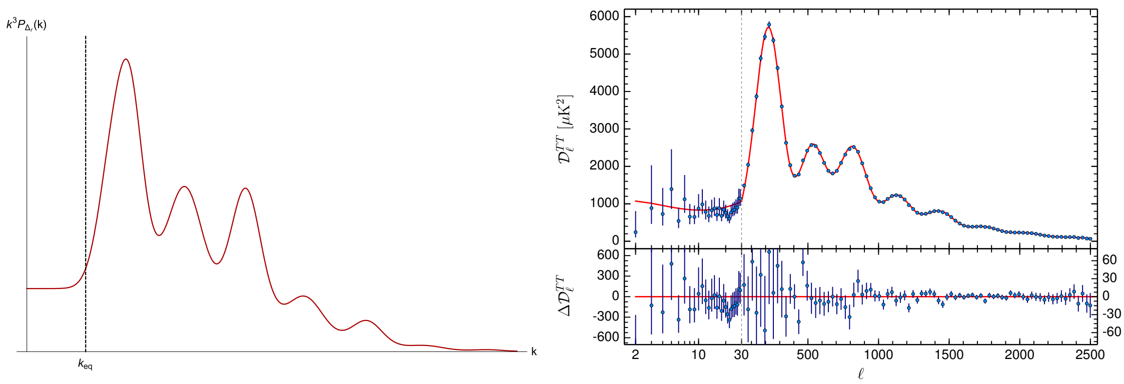


Figure 4.8: Radiation Power Spectrum. Left is our qualitative prediction. Right is observations from Planck2015.

Summary - Linear

In the first part of this chapter the goal was to calculate the matter power spectrum. To do this we first investigated what our initial power spectrum is at some early time when all modes are outside the horizon. By combining (G1) and (G3) we derived a closed equation for the potential,

$$\Phi'' + 2(1+\omega)\mathscr{H}\Phi' - \omega\nabla^2\Phi = 0 \tag{4.61}$$

We solved this on super horizon scales to find $\Phi =$ constant and $\delta = -2\Phi$ which gave us an expression for Φ in terms of ζ

$$\zeta = \frac{5+3\omega}{3+3\omega}\Phi \tag{4.62}$$

so on super horizon scales we know that $\zeta \propto \Phi \propto \delta \propto$ constant and using (P) we found that $\Delta \propto (k\tau)^2\zeta$. Combining this with the predicted power spectrum of ζ from inflation, $P_\zeta \propto k^{n_s-4}$, gives us our initial conditions for all the perturbations.

Now we just need to calculate the subhorizon evolution of the perturbations in both the radiation and matter eras. We find that

- For Φ we use the closed equation for the potential (4.61) to find
  - Radiation Domination $\omega = 1/3 \implies \Phi \propto a^{-2}\cos(\frac{k\tau}{\sqrt{3}})$
  - Matter Domination $\omega = 0 \implies \Phi \propto$ constant.
- For Δ_r we find
  - Radiation Domination (P) $\implies \Delta_r \propto a^2\Phi \propto \cos(\frac{k\tau}{\sqrt{3}})$
  - Matter Domination (C)+(E) $\implies \Delta_r \propto -4\Phi +$ oscillations.
- For Δ_m we find
  - Matter Domination (P) $\implies \Delta_m \propto a\Phi \propto a$
  - Radiation + Matter Domination (C)+(E)+"no fast mode" $\implies$ Mészáros equation

    $$\frac{d^2\delta_m}{dy^2} + \frac{2+3y}{2y(1+y)}\frac{d\delta_m}{dy} - \frac{3}{2y(1+y)}\delta_m = 0 \tag{M}$$

    which gives us the solution $\Delta_m \propto \ln(a)$
- For Δ_b we know that due to Thompson scattering
  - Radiation Domination $\Delta_b \propto \Delta_r \propto \cos(\frac{k\tau}{\sqrt{3}})$
  - Matter Domination (C)+(E) show that $\Delta_b \propto \Delta_m \propto a$ for $\lambda > \lambda_J = c_s\sqrt{\frac{\pi}{G\bar{\rho}}}$ and for $\lambda < \lambda_J$ baryons are pressure supported.

  Finally we showed that the oscillations present in baryon modes which enter the horizon during the radiation era are imprinted in the matter perturbation as Baryon Acoustic Oscillations (BAO)

Then we put this all together to calculate how the inital power spectrum defined by ζ evolved to the matter power spectrum we observe today. The final form was

$$P_m \propto \begin{cases} k^{n_s} & k < k_{eq} \\ \left(1+\ln(\frac{k}{k_{eq}})\right)^2 k^{n_s-4} \times \text{BAO} & k > k_{eq} \end{cases} \tag{4.63}$$

We then did the same for radiation and found the naive prediction (without baryons)

$$P_r \propto k^{n_s-4}\begin{cases} 1 & k < k^* \\ \cos^2(\frac{k\tau^*}{\sqrt{3}}) & k > k^* \end{cases} \tag{4.64}$$

and discussed how it could be improved.

5. Structure Formation - Halos

In the first part of this chapter we have solved the first order perturbation equations for both radiation and matter. However the matter we observe today has undergone non-linear collapse to form bound objects like stars, galaxies and clusters. How can we calculate, say, the distribution of galaxies, or how many we should see at a particular mass scale? In the second part of the chapter we will examine the formation of dark matter halos (stable gravitationally bound objects) which is where we expect galaxies to form. Surprisingly we will see that we can use the linear theory of the previous sections to make reasonable accurate predictions regarding the statistics of non-linear objects like halos.

5.1 Spherical Collapse

First we will examine how a single spherical over density evolves is a homogeneous background. We will consider a flat FRW universe filled only with cold dark matter. We will take a spherical region with radius $\bar{R}$ and compress it into a region with radius $R < \bar{R}$ leaving a gap between the two, see Figure 5.3.

Mass is conserved so we have $\bar{\rho}\bar{R} = \rho R^3$ so $\rho > \bar{\rho}$ and we have created an overdensity. This set up seems odd but it has been specifically designed so we can use Newton's Spherical Mass Theorem to deduce that

1. As the region inside the sphere radius $\bar{R}$ is spherically symmetric as far as the points outside are concerned we can replace our over-density + gap with a uniform density $\bar{\rho}$ and so the background evolution of $\bar{\rho}$ and $\bar{R}$ are unperturbed.
2. As the region outside the sphere radius $\bar{R}$ is spherically symmetric the region inside evolves independently of the background.

Now we are using GR so why should we expect the Newtonian spherical shell theorem to help us? Well firstly when we are considering cold dark matter on subhorizon scales the Newtonian approximation is excellent. Secondly the separation of the evolution of the perturbation and the background continues to hold in GR by Birkhoff's theorem. This tells us that the background will

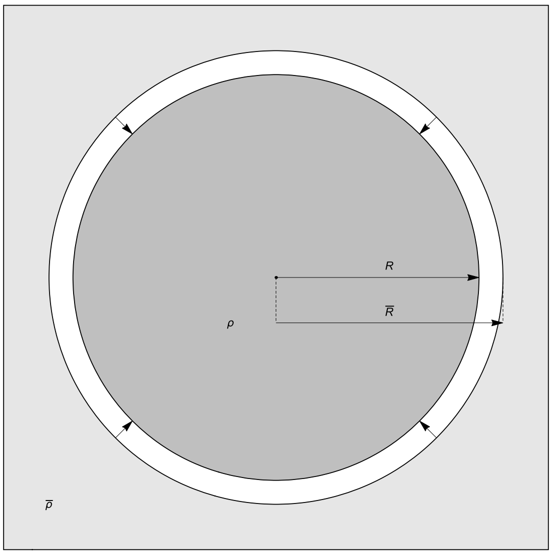


Figure 5.1: Our construction of a spherical over density

evolve just like a flat FRW universe so

$$\bar{\rho} = \bar{\rho}_0 a(t)^{-3} \tag{5.1}$$
$$\bar{R} = \bar{R}_0 a(t) \tag{5.2}$$
$$a(t) = \left(\frac{3}{2}H_0 t\right)^{\frac{2}{3}} \tag{5.3}$$

How does the perturbation evolve? If we set $R \to \infty$ then the solution must be that for a closed universe with density ρ. The spherical shell theorem states that the perturbation can't know anything about matter outside R so the perturbation must obey a scaled down version of the same solution so we have (from Sheet 1, Question 2)

$$\rho = \rho_0 \left(\frac{R_0}{R(t)}\right)^3 \tag{5.4}$$
$$R = R_0 \left[A\left(1 - \cos(\theta)\right)\right], \quad A = \frac{\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \tag{5.5}$$
$$t = B\left(\theta - \sin(\theta)\right), \quad B = \frac{\Omega_{m,0}}{2H_0\left(\Omega_{m,0} - 1\right)^{\frac{3}{2}}} \tag{5.6}$$

where $\theta \in [0, 2\pi)$. Lets expand R and t in θ to see how the overdensity evolves initially. The zeroth order part is

$$R = R_0 \left(A\frac{\theta^2}{2}\right), \quad t = B\frac{\theta^3}{6} \tag{5.7}$$
$$\implies R = R_0 \frac{A}{2}\left(\frac{6t}{B}\right)^{\frac{2}{3}} = R_0 \Omega_{m,0}^{\frac{1}{3}} \left(\frac{3}{2}H_0 t\right)^{\frac{2}{3}} = \bar{R}_0 a(t) = \bar{R} \tag{5.8}$$

which matches the background evolution perfectly. Now if we consider the first order perturbation to this we have

$$R = R_0\left(A\frac{\theta^2}{2}\left[1-\frac{\theta^2}{12}\right]\right), \quad t = B\frac{\theta^3}{6}\left[1-\frac{\theta^2}{20}\right] \tag{5.9}$$

$$\implies R = R_0\frac{A}{2}\left(\frac{6t}{B}\right)^{\frac{2}{3}}\left[1-\frac{1}{20}\left(\frac{6t}{B}\right)^{\frac{2}{3}}+\dots\right] \tag{5.10}$$

$$= \bar{R}_0 a(t)\left[1-\frac{a(t)}{10A}+\dots\right] \tag{5.11}$$

We know mass is conserved so

$$\rho_i R_i^3 = \rho R^3 = \rho_i (1+\delta) R_i^3 (1+\delta_R)^3 \tag{5.12}$$

$$\implies \delta_{linear} \approx -3\delta_R = \frac{3}{20}\left(\frac{6t}{B}\right)^{\frac{2}{3}} = \left(\frac{3}{10AR_0}\right)a(t) \tag{5.13}$$

and so we see that the linear perturbation grows as $a(t)$, matching what we found for matter domination in the previous half of this chapter. Now let us see what linear theory predicts for the overdensity at the key points during the perturbations collapse

- **Turnaround** is when the overdensity stops expanding and begins to collapse. This happens when $\theta = \pi$ so $t = \pi B$ and

$$\delta_{linear} = \frac{3}{20}(6\pi)^{\frac{2}{3}} \approx 1.06 \tag{5.14}$$

- **Collapse** is when the overdensity stops shrinking. This happens when $\theta = 2\pi$ so $t = 2\pi B$ and

$$\delta_{linear} = \frac{3}{20}(12\pi)^{\frac{2}{3}} \approx 1.69 \tag{5.15}$$

This is clearly completely wrong so why are we interested? It will turn out that we will be able to use $\delta = 1.69$ as a critical density at which we expect linear perturbations to have collapsed and formed halos. This allows us to map our linear evolution to the full non-linear case.

Now we will examine what actually happens to these overdensities. First the real overdensity will be

$$1+\delta = \frac{\rho}{\bar{\rho}} = \frac{\rho_0 (R/R_0)^{-3}}{\bar{\rho}_0 a^{-3}} = \frac{\Omega_{m,0}\bar{\rho}_0}{\bar{\rho}_0}\left(\frac{R_0 a}{R}\right)^3 = \Omega_{m,0}\left(\frac{R_0 a}{R}\right)^3 \tag{5.16}$$

And at turnaround we have $a = \left(\frac{3}{2}H_0 t\right)^{2/3}$, $R/R_0 = 2A$, and $t = \pi B$ so

$$1+\delta_{turn} = \Omega_{m,0}\frac{\left(\frac{3}{2}H_0\pi B\right)^2}{(2A)^3} = \frac{9\pi^2}{16} \approx 5.55 \tag{5.17}$$

At collapse we have $R \to 0$ so $\delta \to \infty$ but in reality this will not occur. During collapse small perturbations will grow which convert the KE of collapse into random motion of the dark matter particles. This random motion will eventually balance the gravitational collapse and the system will reach a stable equilibrium forming a halo with fixed radius and mass. This process is called **Virialisation**. We can calculate when this will occur by using the virial theorem (see Wikipedia for a proof):

Virial Theorem

$$\text{Potential Energy}\,(V) = -2 \times \text{Kinetic Energy}\,(T) \tag{V}$$

So when does virialisation happen? We know that at turnaround the kinetic energy is zero so all the energy is potential energy

$$E = V_{turn} \tag{5.18}$$

And as energy is conserved we find that at virialisation, when $V = -2T$,

$$E = V_{vir} + T_{vir} = \frac{1}{2} V_{vir} \tag{5.19}$$

So we know that at virialisation the potential energy has doubled from turnaround. This requires

$$V_{vir} = 2V_{turn} \tag{5.20}$$

$$R_{vir} = R_{turn}/2 \tag{5.21}$$

$$\rho_{vir} = 8\rho_{turn} \tag{5.22}$$

so the density has increased by a factor 8 from turnaround so $\rho = 8 \times 5.55 \times \bar{\rho}_{turn}$. We now just need to relate $\bar{\rho}_{turn}$ to $\bar{\rho}_{vir}$. From turnaround $t = \pi B$ to virialisation $t = 2\pi B$ the background density has decreased by a factor

$$\bar{\rho}_{turn} = \left(\frac{a_{vir}}{a_{turn}}\right)^3 \bar{\rho}_{vir} \tag{5.23}$$

$$= \left(\frac{t_{vir}}{t_{turn}}\right)^2 \bar{\rho}_{vir} \tag{5.24}$$

$$= 4\bar{\rho}_{vir} \tag{5.25}$$

So putting it all together we have

$$\rho_{vir} = 5.55 \times 8 \times 4 \times \bar{\rho}_{vir} \tag{5.26}$$

$$\approx 178\bar{\rho}_{vir} \tag{5.27}$$

This gives us the clear intuition from linear theory.

Whenever the linear overdensity exceeds $\delta = 1.69$ we should expect that a halo has formed with a density roughly 200 times the background see Figure 5.2

5.2 Filtering

So in the previous section we found that if we consider spherical overdensities then whenever the linear perturbation exceeds 1.69 then a halo should form with density $\rho \approx 200 \times \bar{\rho}$. To make use of this insight we will first smooth the field δ on some length scale R so that we can associate a spherical overdensity with radius R with each point. This will also allow us to easily associate a mass ($\propto \bar{\rho}R^3$) to the density perturbation at each point.

We will filter the density perturbation via convolution with a window function, W, so

$$\delta_R(t, \mathbf{x}) = \int \delta(t, \mathbf{x}') W(|\mathbf{x} - \mathbf{x}'|, R) d^3\mathbf{x}'. \tag{5.28}$$

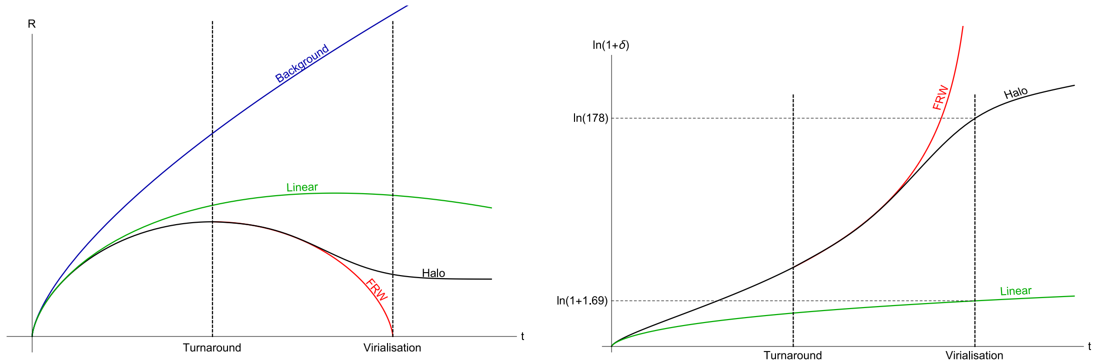


Figure 5.2: Evolution of the spherical over density

Or in Fourier space, where the convolution becomes a product,

$$\delta_R(t,k) = \delta(t,k)\widetilde{W}(kR), \quad \widetilde{W}(kR) \equiv \int W(\mathbf{x},R)e^{-i\mathbf{k}\cdot\mathbf{x}}d^3\mathbf{x}. \tag{5.29}$$

Popular Window Functions:

- **Top Hat**

$$W_{TH}(r,R) = \begin{cases} \frac{3}{4\pi R^3} & r \leq R \\ 0 & r > R \end{cases}, \quad \widetilde{W}_{TH}(kR) = \frac{3}{(kR)^3}\left(\sin(kR) - (kR)\cos(kR)\right) \tag{5.30}$$

- **Gaussian**

$$W_G(r,R) = \frac{1}{\left(\sqrt{2\pi}R\right)^3}\exp\left(-\frac{r^2}{2R^2}\right), \quad \widetilde{W}_G(kR) = \exp\left(-\frac{(kR)^2}{2}\right) \tag{5.31}$$

- ***k*-cut**

$$W_{k\text{-cut}}(r,R) = \frac{1}{2\pi^2 r^3}\left(\sin\left(\frac{r}{R}\right) - \left(\frac{r}{R}\right)\cos\left(\frac{r}{R}\right)\right), \quad \widetilde{W}_{k\text{-cut}}(kR) = \begin{cases} 1 & k \leq \frac{1}{R} \\ 0 & k > \frac{1}{R} \end{cases} \tag{5.32}$$

As we are smoothing the linear field we will assume it is Gaussian (and large non-Gaussianities in the inflationary spectrum are tightly constrained by Planck2015). This means that the field is completely described by its mean and variance.

$$\bar{\delta}_R \equiv \langle \delta_R \rangle = 0 \tag{5.33}$$

$$\sigma_R^2 \equiv \langle \delta_R^2 \rangle = \frac{1}{2\pi^2}\int P(k)\widetilde{W}^2(kR)k^2 dk \tag{5.34}$$

From this we can define an important cosmological parameter

σ_8

$$\sigma_8 \equiv \sqrt{\frac{1}{2\pi^2}\int P_{linear}(k)\widetilde{W}_{TH}^2(kR)k^2 dk}, \quad \text{where} \quad R = 8\mathrm{h}^{-1}\mathrm{Mpc} \tag{S$_8$}$$

σ_8 is the variance of the linear density field on 8Mpc scales and it acts to normalise the matter power spectrum. Observations put $\sigma_8 \approx 0.8$. If σ_8 were larger then we would have larger fluctuations and structure would have formed earlier, conversely if σ_8 were smaller structure would have formed later. Now how does σ_R depend on R? We know from the previous chapter that

$$P_{linear} \propto \begin{cases} k^{n_s} & k < k_{eq} \\ k^{n_s-4} & k > k_{eq} \end{cases} \equiv A^2(t) k^{n_{eff}} \tag{5.35}$$

where we have defined an effective spectral index $n_{eff} = n_{eff}(k)$ and $A(t) \propto a(t)$. On the scales we are interested in for large scale structure we usually have $k \geq k_{eq}$ and $n_{eff} \approx -3$ and reasonably slowly varying so this definition make more sense than first appears. This gives us

$$\sigma_R^2(t) = \frac{A^2(t)}{2\pi^2} \int k^{n_{eff}+2} \widetilde{W}^2(kR)\, dk \tag{5.36}$$

$$\approx \frac{A^2(t)}{2\pi^2} \frac{1}{R^{n_{eff}+3}} \int y^{n_{eff}+2} \widetilde{W}^2(y)\, dy, \quad y \equiv kR \tag{5.37}$$

Where we have evaluated n_{eff} for $k = \frac{2\pi}{R}$ and assumed it changes slowly. Also in the range of interest $0.8\text{Mpc} < R < 40\text{Mpc}$ the integral over y is approximately constant (it changes by a factor of ≈ 2 which is irrelevant for our calculations). So we have

$$\sigma_R(t) \propto a(t) R^{-\left(\frac{n_{eff}+3}{2}\right)} \tag{5.38}$$

It will be more useful to us if we convert this to mass where

$$\sigma_R = \sigma_M \quad \text{for} \quad M(R) = V_R \int \rho\, W(\mathbf{x}, R)\, d^3\mathbf{x} \tag{5.39}$$

and V_R is the volume of the window function W

Exercise: Show

$$\sigma_M^2 = \left\langle \left(\frac{M - \bar{M}}{\bar{M}} \right)^2 \right\rangle \tag{5.40}$$

Then we have that the dependence of σ_M on M is

$$\sigma_M \propto a(t) M^{-\frac{n_{eff}+3}{6}} \tag{5.41}$$

5.3 Extended Press-Schechter Mass Function

We can now use the ideas in the previous two sections to calculate the mass function, which is the number density of halos in a given mass range. We know that if the density at a point in linear theory is larger than 1.69 then we will form a halo. We can associate a mass with that point by filtering with a window function. The question is now what smoothing scale should we use to find the mass of the halo?

Suppose we smooth the density field on two different scales, R_1 and R_2, and the density perturbation is still larger than 1.69 for both at a given point, $\delta_{R_1}, \delta_{R_2} > 1.69$. Then we expect a halo to form and have a choice of which mass to associate with it, M_{R_1} or M_{R_2}, which should we

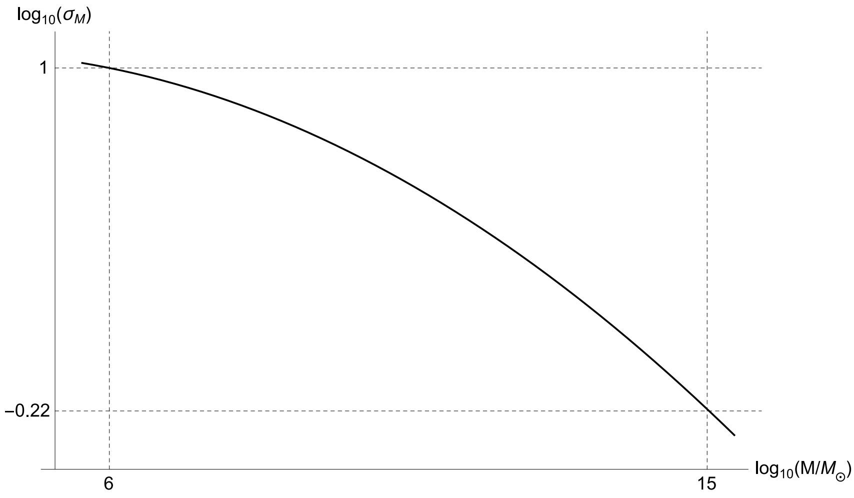


Figure 5.3: Dependence of σ_M on mass. We see that in 9 decades of mass σ_M only changes by 1 decade and is close to linear justifying our use of n_{eff}

chose? It should be clear that we should choose the larger of the two values as we would expect the collection of smaller halos close together to merge and form one large halo, see Figure 5.4. This leads to the following insight:

Extended Press-Schechter Assumption

We should associate a halo with a given point in space (and time) by choosing the largest possible smoothing scale, R, for which $\delta_R > 1.69$. The mass of this halo will then be M_R

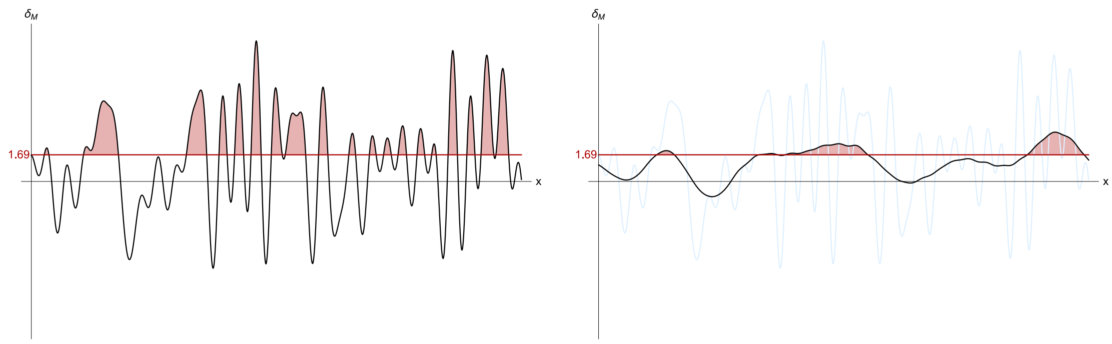


Figure 5.4: Smoothing the density field with two length scales, a short on on the left and a long one on the right. We see that the density field could either be interpreted as corresponding to a large number of small halos when we smooth with a small R as on the left or as having three large halos when we smooth with a large R as on the right. The interpretation on the right is the correct one as we would expect clusters of small halos like those on the left to merge.

Now we need to consider how to find the largest possible smoothing scale for which the density exceeds the critical value. A simple approach might be to simply extend the smoothing scale until the density at a point falls below the critical value and call this the maximum scale, this however neglects the possibility that there is some even larger scale for which the smoothed density once again rises above the critical value (this is is known variously as the cloud/peak/halo in cloud/peak/halo problem). Instead we must start by smoothing with $R = \infty$ and decrease it until $\delta_S = \delta_C$.

To do so it will help to think about things in a slightly different way. First we will use $S = \sigma_R^2$ rather than R to denote our smoothing scale. From Eq. 5.38 that $S = \sigma_R^2$ and R are inversely related and that the relationship is monotonic so this relabelling is allowed. Now $S = 0$ corresponds to $R \to \infty$ where $\rho = \bar{\rho}$ (by definition of being the background) so $\delta_S = 0$. Then as we increase S (decrease R) δ_S will perform a random walk. We will specialise to the case where we use a k-space cut for our window function as this guarantees that independent information is removed as we increase S so the random walk is Markovian (what it does at each step does not depend on the past), see Figure 5.5. Now the problem of finding the largest possible scale for which the smoothed density perturbation exceeds the critical value is the problem of finding the first crossing of a barrier, $\delta_C = 1.69$, for a Markovian random walk with S playing the role of time.

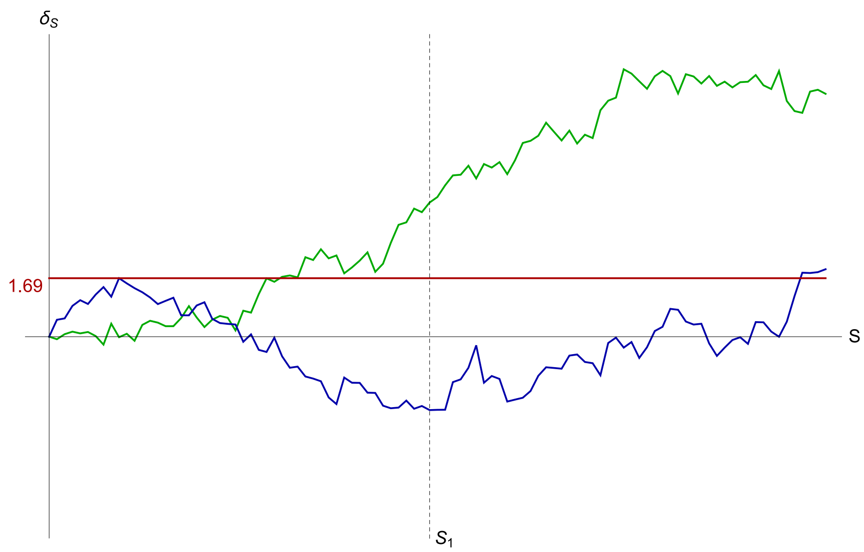


Figure 5.5: Example of two possible random walks. The first (in green) first-crosses before S_1 and then remains above for all S. The second (in blue) first-crosses even earlier (so corresponds to larger halo) but then dips back below and at S_1 looks like an underdensity.

So how do we calculate the probably that a trajectory has achieved first crossing before a given S_1 (i.e. the probability that a halo with mass larger than M_{S_1} has formed)? We know that all trajectories that are above the critical line at a given S_1 must have crossed before that time. So if we denote S_X as the S for which a trajectory first crosses the barrier then we know that (as the random walks are Gaussian distributed),

$$P(S_X < S_1) > P(\delta_{S_1} > \delta_c) = \frac{1}{\sqrt{2\pi}\sigma_M} \int_{\delta_c}^{\infty} \exp\left(-\frac{x^2}{2\sigma_M^2}\right) dx \tag{5.42}$$

$$= \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_M}\right) \tag{5.43}$$

$$= \frac{1}{2}\mathrm{erfc}\left(\frac{\nu}{\sqrt{2}}\right) \tag{5.44}$$

But how do we account for trajectories which have crossed **before** S_1 but are below the critical line **at** S_1? Here we use a neat trick using the fact that the random walk is Markovian. This means that as a trajectory touches the barrier the probability of it going either up or down in the next step is equal. So for every trajectory that crosses before S_1, but is under at S_1, there is an image trajectory, reflected in the critical line, above S_1, which is equally likely, see Figure 5.6. This means that to include the all trajectories which have dipped back below the line all we need to do it multiply by 2!

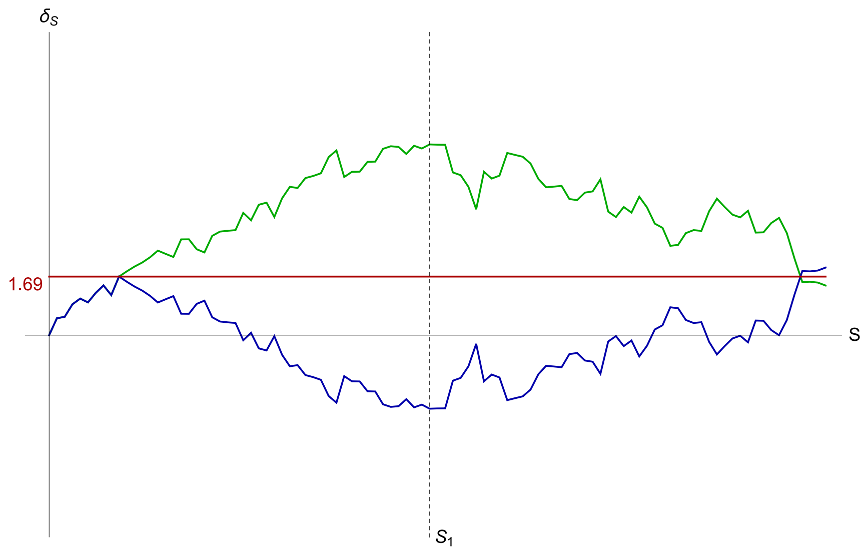


Figure 5.6: A random walk which crosses before S_1 but is below at S_1 (in Blue) and its mirror trajectory (in Green) which is above the critical value at S_1.

$$P(S_X < S_1) = 2P(\delta_{S_1} > \delta_c) = \operatorname{erfc}\left(\frac{\nu}{\sqrt{2}}\right) \tag{5.45}$$

This is the probability for of forming a halo with mass $> M_{S_1}$, but what we really want is the mass function, the mean number density of halos mass M per unit of comoving volume. The volume associated with a halo of mass M (the region of space that the mass came from) is

$$V_M = \frac{M}{\bar{\rho}} \tag{5.46}$$

and the fraction of halos in the mass range $[M, M+\Delta M]$ is

$$-\frac{dP}{dM}\Delta M \tag{5.47}$$

So putting it together we have the mass function

$$\frac{d\bar{n}_h}{dM} = -\frac{1}{V_M}\frac{dP}{dM} = \frac{-\bar{\rho}}{M}\frac{dP}{d\sigma_M}\frac{d\sigma_M}{dM} \tag{5.48}$$

which we can calculate as

Press-Schechter Mass Function

$$\frac{d\bar{n}_h}{dM} = -\sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M\sigma_M}\frac{d\sigma_M}{dM}\nu\exp\left(-\frac{\nu^2}{2}\right) \tag{5.49}$$

We know from the previous section that

$$\frac{\nu}{\sqrt{2}} = \frac{\delta_c}{\sqrt{2}\sigma_M} = B(t)M^{\frac{3+n_{eff}}{6}} \tag{5.50}$$

so if we define $\gamma \equiv 1 + n_{eff}/3$ then we can write the Press-Schechter mass function as

$$\frac{d\bar{n}_h}{dM} = \frac{\gamma}{\sqrt{\pi}} \frac{\bar{\rho}}{M^2} B(t) M^{\frac{\gamma}{2}} \exp\left(-B^2(t)M^{\gamma}\right) \tag{5.51}$$

which has the following behaviour

- Small $M \to$ Power law
- Large $M \to$ Exponential cut off

When we compare this to data we over-predict low mass halos by about a factor of 2 and under predict high mass halos by about a factor of 10 which is remarkable considering how naive the derivation is, see Figure 5.7.

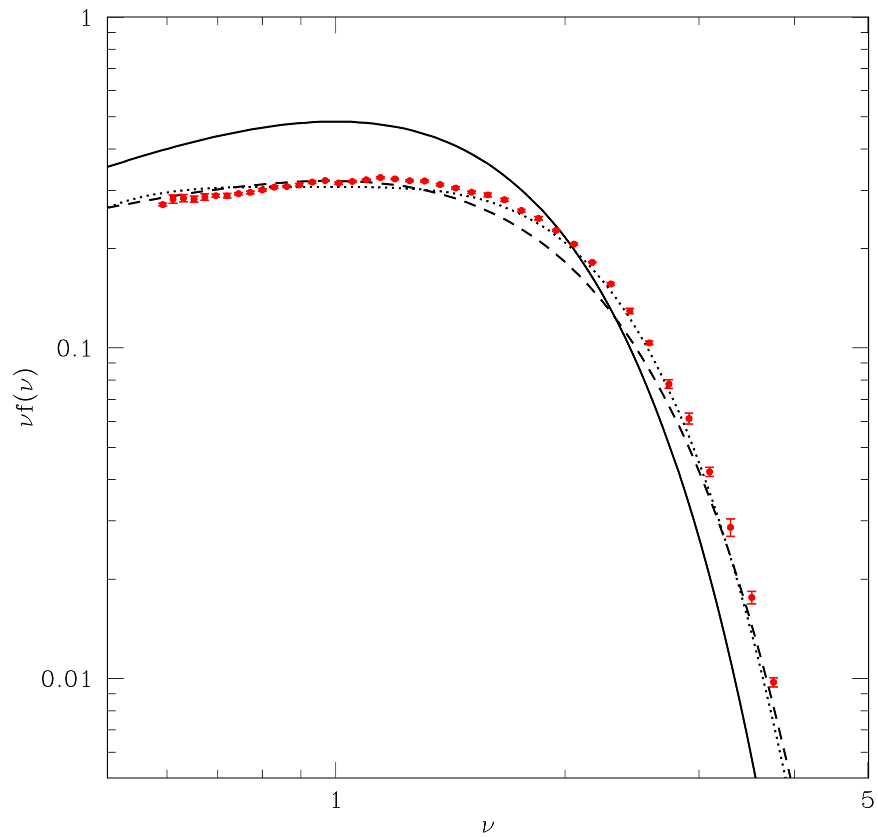


Figure 5.7: Press-Schechter versus results from N-body simulations. PS theory is the black solid line. The two dashed and dotted lines other more advanced methods and red is data from simulations. $\nu f(\nu) \propto \frac{M^2}{\bar{\rho}} \frac{d\bar{n}_h}{dM}$ which makes the two limiting behaviours easier to see.

5.4 Linear Bias

We have found a way to use linear theory to calculate the number density for halos so we know how many we should observe in a volume for a given mass range. The next question is how will they be distributed in that volume, will they be scattered randomly or clumped together? We can use PS theory to calculate how correlation functions for the linear field are related to correlation functions for the halos. This ratio of the two is called the **bias**.

First what statistics do we use to measure how objects are distributed in space? If we define the Halo over density

$$\delta_h = \frac{n_h - \bar{n}_h}{\bar{n}_h} \tag{5.52}$$

Then we can define the halo-halo 2-point correlation function

$$\xi_{hh}(r) = \left\langle \delta_h(\mathbf{x})\delta_h(\mathbf{x}')\right\rangle = \frac{\langle n_h(\mathbf{x})n_h(\mathbf{x}')\rangle}{\bar{n}_h^2} - 1 \tag{5.53}$$

This tells us if we have a halo how likely are we to find another at a distance $r = \mathbf{x} - \mathbf{x}'$ (its Fourier transform is the power spectrum we met earlier, the two contain the same information just in a different form)

Now the trick to using PS to calculate the bias is to make what is called the Peak-Background split. This is when we break up the density perturbation into short wavelength "peaks" which are what will go on to form halos, and longer wavelength "background" these modes are too big to form halos so just act as a modulation to the background.

$$\delta = \delta_h + \delta_b \tag{5.54}$$

where δ_h are the short wavelength modes or "peaks" which will form halos and δ_b are the long wavelength modes which will modulate the background which will be linear, see Figure 5.8. Now

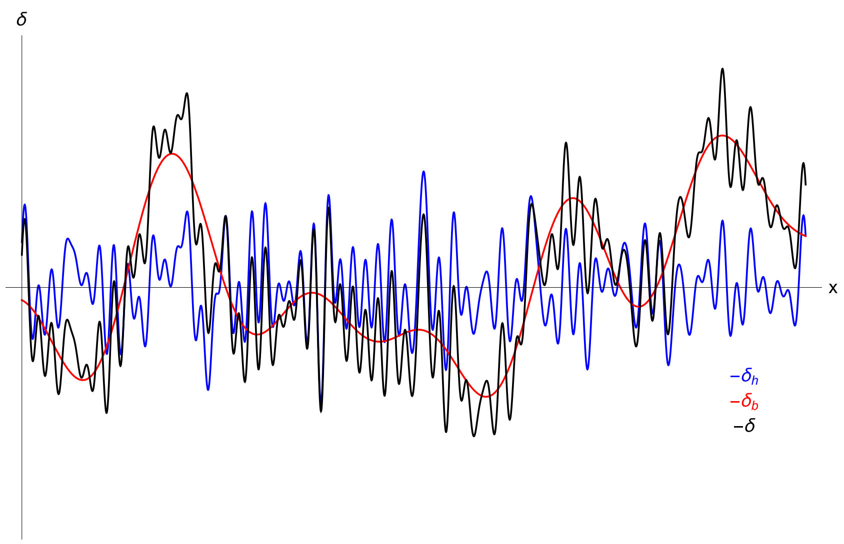


Figure 5.8: The Peak-Background split. Red is the long wavelength background, blue is the short wavelength peaks and black is their sum.

we need to determine the dependence of n_h on δ_b. From PS we know that $\frac{dn_h}{dM}$ depends on δ_b in two ways: The first is through $\bar{\rho}$. We expect that locally the peaks will see a background density,

$$\widetilde{\rho} \approx \bar{\rho}\,(1+\delta_b)\,. \tag{5.55}$$

The second is through δ_c as the critical density will seem lower in regions where δ_b is high so

$$\widetilde{\delta_c} \approx \delta_c - \delta_b \tag{5.56}$$

So if we expand $\frac{dn_h}{dM}$ around δ_b we have

$$\frac{dn_h}{dM}(\delta_b) \approx \frac{d\bar{n}_h}{dM} + \left[\frac{\partial}{\partial\widetilde{\delta_c}}\left(\frac{d\bar{n}_h}{dM}\right)\frac{d\widetilde{\delta_c}}{d\delta_b} + \frac{\partial}{\partial\widetilde{\rho}}\left(\frac{d\bar{n}_h}{dM}\right)\frac{d\widetilde{\rho}}{d\delta_b}\right]\delta_b + \dots \tag{5.57}$$

$$= \frac{d\bar{n}_h}{dM}\left[1 + \left(\frac{\nu^2-1}{\nu\sigma_M} + 1\right)\delta_b + \dots\right] \tag{5.58}$$

We can use this then to show

$$\delta_h = \frac{n_h - \bar{n}_h}{\bar{n}_h} = \frac{\frac{dn_h}{dM} - \frac{d\bar{n}_h}{dM}}{\frac{d\bar{n}_h}{dM}} = \left(\frac{\nu^2-1}{\delta_c} + 1\right)\delta_b = b(M)\delta_b \tag{5.59}$$

Where we have defined the linear bias as

$$b(M) \equiv \frac{\nu^2-1}{\delta_c} + 1 \tag{5.60}$$

see Figure 5.9. Now we can relate the halo-halo 2-point correlator to the 2-point correlator for the linear field by

$$\xi_{hh}(r,M) = b^2(M)\xi_{linear}(r) \tag{5.61}$$

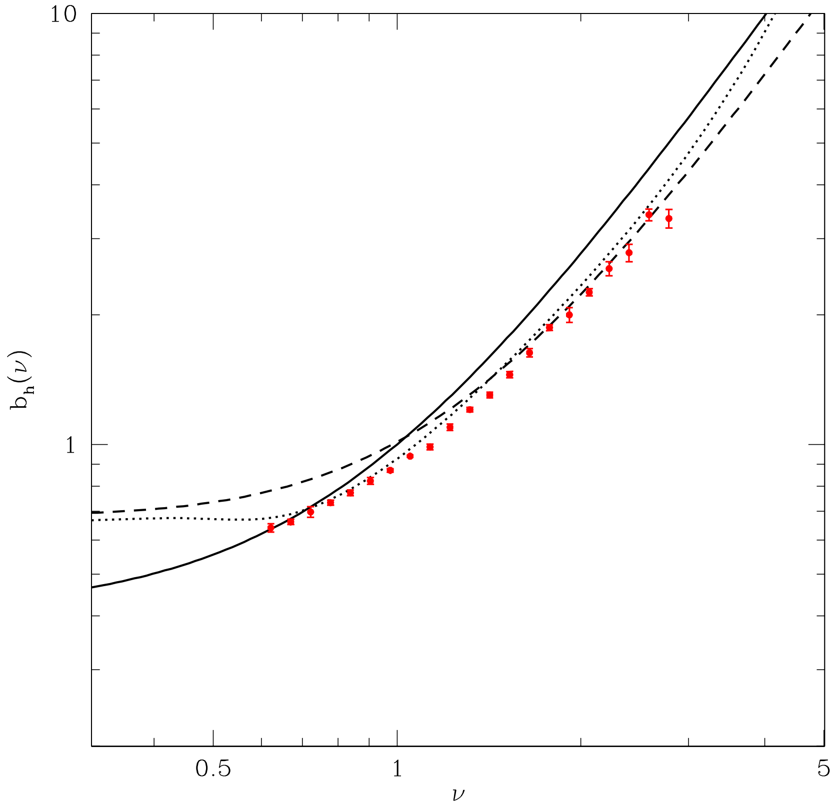


Figure 5.9: Bias for PS (Black) and improved methods versus data from simulations

Summary - Halos

In the second part of this chapter we considered how we can use the linear predictions of the first half to calculate statistics of non-linear dark matter halos. We began by considering a toy model of collapse, that of a spherical overdensity in flat background. We found that as the spherical perturbation and the background decouple the perturbation follows the solution for an overdense universe,

$$R = R_0 \left[A \left(1 - \cos(\theta) \right) \right], \quad A = \frac{\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \tag{5.62}$$

$$t = B \left(\theta - \sin(\theta) \right), \quad B = \frac{\Omega_{m,0}}{2H_0 \left(\Omega_{m,0} - 1 \right)^{\frac{3}{2}}} \tag{5.63}$$

we linearised this solution to find

$$\delta_{linear} = \frac{3}{20} \left(\frac{6t}{B} \right)^{\frac{2}{3}} \propto a \tag{5.64}$$

as we found from the linearised equations. We then calculated the full non-linear solution, introducing the idea of virialisation, and found that we formed a stable collapsed object when $R_{vir} = \frac{1}{2} R_{turn}$. With this we mapped the linear solution to the non-linear one to find:

	δ_{linear}	$1+\delta$
Turnaround	1.06	5.55
Virialisation	1.69	178

This gives us the insight that when the linear density contrast exceeds 1.69 we should expect a halo to form of mass $\approx 200\bar{\rho}$.

We then introduced the idea of filtering the density perturbation so we can associate perturbations with a mass scale (M)

$$\delta_M(t,\mathbf{x}) = \int \delta(t,\mathbf{x}') W(|\mathbf{x} - \mathbf{x}'|, R) d^3\mathbf{x}', \quad M(R) = V_R \int \rho \, W(\mathbf{x}, R) \, d^3\mathbf{x} \tag{5.65}$$

and calculated the variance of the smoothed field to be,

$$\sigma_R^2 \equiv \left\langle \delta_R^2 \right\rangle = \frac{1}{2\pi^2} \int P(k) \, \widetilde{W}^2(kR) \, k^2 dk \tag{5.66}$$

which led to the definition of σ_8. Next we introduced the effective spectral index n_{eff} for the matter power spectrum and found that

$$\sigma_M \propto a(t) M^{-\frac{n_{eff}+3}{6}} \tag{5.67}$$

where for most observational scales $n_{eff} \approx n_s - 4$

We then use the insights of the previous two sections plus the assumption that the fraction of mass in halos above a mass scale M is just the probability that $\delta_M > 1.69$ to derive the Press-Schechter mass function

$$\frac{d\bar{n}_h}{dM} = -\sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M\sigma_M} \frac{d\sigma_M}{dM} \nu \exp\left(-\frac{\nu^2}{2} \right) \tag{5.68}$$

which we showed has the following behaviour:

- small M $\implies \frac{d\bar{n}_h}{dM} \propto M^{\frac{\gamma}{2}-2}$, so a power law

- large M $\implies$ exponential cut-off.

Finally we investigated how the halos were distributed and found they are biased relative to the linear perturbation. We defined the halo overdensity

$$\delta_h \equiv \frac{n_h - \bar{n}_h}{\bar{n}_h} \tag{5.69}$$

and used the peak background split to show that

$$\frac{dn_h}{dM}(\delta_b) = \frac{d\bar{n}_h}{dM}\left[1 + \left(\frac{\nu^2 - 1}{\nu\sigma_M} + 1\right)\delta\right] \tag{5.70}$$

so $\delta_h = b(M)\delta$ where

$$b(M) = \frac{\nu^2 - 1}{\delta_c} + 1 \tag{5.71}$$

is the linear bias.